University of Tartu
Faculty of Science and Technology
Institute of Physics

Karl Ehatäht

Search for $t\bar{t}h$ production in $3\ell + 1\tau_h$ in final state

Master's Thesis

Supervisors:
PhD Christian Veelken
PhD Luca Marzola

Tartu 2017

Title
Search for $t\bar{t}h$ production in $3\ell + 1\tau_h$ in final state

Abstract
This thesis presents a measurement of the cross section for the associated production of a Higgs boson with a pair of top quarks in events containing 3 electrons or muons and one hadronically decaying τ. The analyzed data, collected with the CMS detector in 2016, amounts to an integrated luminosity of $35.9\,\text{fb}^{-1}$, and has been recorded in proton-proton collisions at c.o.m energy of $\sqrt{s} = 13\,\text{TeV}$ with bunch spacing of 25 ns.

The signal strength ($\mu = \sigma/\sigma_{\text{SM}}$) has been measured to $\mu = 1.22^{+1.34}_{-1.00}$, while the expected is $\mu = 1.00^{+1.41}_{-1.06}$. This result corresponds to cross section measurement of $\sigma(t\bar{t}h) = 618.7^{+679.5}_{-507.1}\,\text{fb}$, whereas the SM predicts $\sigma_{\text{SM}}(t\bar{t}h) = 507.1\,\text{fb}$. The results are compatible with the SM predictions, but relatively large uncertainties and low significance of overall $h \to \tau\tau$ decay channel leave room for improvements.

Keywords
top physics, Higgs physics, CMS

CERCS code
P210 (elementary particle physics, quantum field theory)

Pealkiri
Protsessi $t\bar{t}h$ uurimine $3\ell + 1\tau_h$ lõppolekuga

Kokkuvõte
Antud töö esitab mõjuristlõike mõõtmistulemuse analüüsi protsessi jaoks, kus Higgsi boson tekib koos top-kvargi paariga, ning mille lõppolekusse kuulub kolm leptonit ja üks hadroniline τ laguprodukt (tähistatuna τ_h). Analüüsitud andmed on kogutud CMSi detektoriga 2016. aastal prooton-prooton põrgetest massikeskme energial $\sqrt{s} = 13\,\text{TeV}$ ja prootonite kimbu vahekaugusel 25 ns. Andmete hulk vastab integreeritud luminositeedile $35.9\,\text{fb}^{-1}$.

Signaali tugevuseks (ehk suhe mõõdetud ja Standardmudeli ennustatud mõjuristlõigete vahel, $\mu = \sigma/\sigma_{SM}$) mõõdeti $\mu = 1.22^{+1.34}_{-1.00}$ ja ooteväärtuseks hinnati $\mu = 1.00^{+1.41}_{-1.06}$. Antud mõõtetulemus vastab mõjuristlõikele $\sigma(t\bar{t}h) = 618.7^{+679.5}_{-507.1}\,\text{fb}$, samas kui Standardmudel ennustab $\sigma_{\text{SM}}(t\bar{t}h) = 507.1\,\text{fb}$. Tulemused langevad kokku Standardmudeli ennustusega, ent suured määramatused ja üldiselt madal statistiline tähtsus $h \to \tau\tau$ lagunemiskanalis jätab palju arenguruumi järgnevateks analüüsideks.

Märksõnad
top kvark, Higgsi boson, CMS

CERCS kood
P210 (elementaarosakeste füüsika, kvantväljade teooria)

Contents

1 Introduction **1**

2 The Standard Model & the Higgs boson **3**
2.1 The Standard Model 3
2.2 The Higgs boson 6
2.3 Higgs phenomenology 10

3 Experimental setup **16**
3.1 The LHC 16
3.2 The CMS detector 19
3.2.1 Silicon tracker 20
3.2.2 Calorimeters 20
3.2.3 Muon detector 22
3.2.4 Triggers 23

4 The analysis **26**
4.1 Data and MC samples 26
4.2 Object reconstruction and identification 27
4.2.1 Electrons 27
4.2.2 Muons 28
4.2.3 Jets 29
4.2.4 Hadronic τ decays 30
4.2.5 Event-level variables 31
4.3 Event selection 31
4.3.1 Data-to-MC corrections 33
4.3.2 Background estimation 34
4.4 Signal extraction 35
4.4.1 Systematic uncertainties 36
4.4.2 Final discriminant 37
4.4.3 Matrix element method 40
4.4.4 Results 46

5 Summary **47**

Bibliography **48**

Abbreviations **55**

Acknowledgments **57**

Appendix A The Standard Model **58**
A.1 The Higgs doublet 58
A.2 Gauge-kinetic term of the Higgs field 59
A.3 Potential term of the Higgs field 61

A.4 Quark sector . . . 61
A.5 Decay widths and branching ratios . . . 62

Appendix B CMS coordinate system and kinematics . . . 64
B.1 Coordinate system . . . 64
B.2 Missing transverse energy . . . 65
B.3 Isolation . . . 66
B.4 Vertexing . . . 67

Appendix C Analysis details . . . 69
C.1 Fake factor method . . . 69
C.2 Statistical treatment . . . 73

Appendix D Tables . . . 76

1. Introduction

The Standard Model (SM) is still regarded as one of the most successful theories in modern physics due to its realm and predictive power. It describes three of the four fundamental forces – strong, weak and electromagnetic – mediated by and acting on elementary particles such as gluons, weak bosons, photons, quarks, leptons and neutrinos. The Higgs boson, theorized about fifty years ago, was discovered in 2012 by the ATLAS and CMS experiments [1, 2]. With the recent advances in high energy physics (HEP) and continuous upgrades of the LHC detectors, it is now possible to access the Higgs properties via precision studies. One of the most intriguing aspects yet to test is the coupling of the Higgs boson to the top quark, the heaviest known elementary particle to date. This thesis presents a state of the art analysis of the interaction by searching for $t\bar{t}h$ events in $h \to \tau\tau$ decay channel with the focus on $3\ell + 1\tau_h$ signature. Our goal is to measure the cross section of the signal process.

The thesis is divided into three chapters, each supported by appendices that give a more detailed overview of the topics that are partially or insufficiently explained in the available university courses. The first chapter describes the mathematical foundations of the SM and motivates the necessity of a Higgs boson on experimental and theoretical grounds. This is followed by section 2.3 about Higgs phenomenology, which explores the dominant Higgs production and decay mechanisms, with emphasis on $t\bar{t}h$ production and $h \to \tau\tau$ decay mode. The importance of Higgs-to-top coupling for current physics, as well as the novelty aspect of $3\ell + 1\tau_h$ signature are also explained there.

The second chapter is devoted to a description of the experimental apparatus, in particular the LHC ring (section 3.1) and the CMS detector (section 3.2). All of the analyzed data used in the present work have been recorded with the CMS detector, the physical parameters of which define and constrain the analysis strategy. Each of the subdetectors, their working principles and design-driven specifications are documented. The CMS collaboration has developed sophisticated trigger systems to handle high collision rates, which are also outlined in the text.

The third chapter presents the analysis, first by listing relevant Monte Carlo (MC) simulations and datasets obtained during 2016 data-taking period. The next section (4.2) is dedicated to object reconstruction, i.e. an algorithmic framework which translates raw information to high-level physics objects representing either leptons, jets or hadronic τ decay products (denoted by τ_h). The third section (4.3) provides a list of conditions imposed on the reconstructed objects, with the aim to select signal events and reject background as much as possible. Imperfections of MC simulations w.r.t data are corrected by appropriately weighing the selected MC events. Special attention is paid to model backgrounds due to the misidentification of non-prompt leptons as prompt ones. Finally, systematic uncertainties, signal extraction method and final results are given in section 4.4.

It is important to stress that the analysis presented here is not my individual work, but a joint effort between my home institute, NICPB, and Cornell, LLR and Notre Dame universities, which all independently implemented their versions of the analysis. My contribution to implementation of the analysis developed by the NICPB team includes: synchronization between different institutes participating in the analysis; preparation of data and MC samples for processing; development, troubleshooting and solving technical issues in our framework; implementation of the matrix element method (MEM) for $t\bar{t}h$

signal and $t\bar{t}Z$ background processes that correspond to $3\ell + 1\tau_h$ final state. The analysis is implemented mostly in C++ and Python programming languages, and is hosted in a public repository [3]. My implementation of the MEM is available at [4]. The analysis of $t\bar{t}h$ production with final states involving τ_h, including the work I am about to present, has already been published in [5].

2. The Standard Model & the Higgs boson

The Standard Model (SM) is the best description of particle physics to date, as its predictive power and mathematical consistency have been repeatedly proven by the measurements of its parameters in particle accelerators. The first groundwork of the SM was laid out by Glashow, Salam and Weinberg with the electroweak unification in the 1960s [6–8], and with the development of QCD by Fritzsch, Gell-Mann, Gross, Wilczek and Zweig (among many others) in the following decade [9–12]. The SM explains $\sim 5\%$ of the matter-energy content in the Universe and three fundamental forces: the strong, weak and electromagnetic interactions. The elements that are responsible for the remaining 95% of the energy budget of the Universe, as well as the fourth fundamental force (gravity) do not find a description within this theory.

The matter particles (fermions) span three generations of quarks and leptons. The force carriers (gauge bosons) are gluons, the $W^{\pm}$ and Z bosons, and photons, which mediate strong, weak and electromagnetic force, respectively. The SM also includes one scalar particle, the Higgs boson, which is responsible for explaining the masses of quarks, leptons and weak interaction gauge bosons through the spontaneous symmetry breaking (SSB) implemented by Higgs mechanism [13–18]. The existence of this particle has been experimentally verified in 2012 by the ATLAS and CMS collaborations [1, 2], whose combined analyses estimated Higgs mass to be $m_h = 125.09\,\text{GeV}$ [19].

The following discussion about the theoretical settings is mainly adapted from [20–23], numerical values are taken from [24]. The first section (2.1) gives a brief overview of the SM in absence of the Higgs boson; the second section (2.2) extends the framework by including the Higgs boson and examines its consequences; the third section (2.3) explores Higgs decay and production mechanisms. Einstein summation convention, i.e. summing over repeated indices is assumed throughout this text.

2.1. The Standard Model

The SM is encapsulated in a Lagrangian density, which omitting mass terms can be written as

$$\begin{aligned}\mathcal{L}_{\text{SM,massless}} = -\frac{1}{4}\left(G^a_{\mu\nu}G^{\mu\nu}_a + W^a_{\mu\nu}W^{\mu\nu}_a + B_{\mu\nu}B^{\mu\nu}\right) + i\bar{E}_L \not{D} E_L + i\bar{E}_R \not{D} E_R + \\ + i\bar{Q}_L \not{D} Q_L + i\bar{u}_R \not{D} u_R + i\bar{d}_R \not{D} d_R\,,\end{aligned} \tag{2.1}$$

where $\not{D} = \gamma^\mu D_\mu$ is the slashed covariant derivative and γ^μ denotes the Dirac matrices. The first three terms are the gauge field strength tensors of gluon fields $\left\{G^i_\mu\right\}_{i=1}^{8}$, electroweak (EW) fields $\left\{W^i_\mu\right\}_{i=1}^{3}$ and field B_μ, respectively[1]. A particular field strength tensor $A_{\mu\nu} = T_i \mathcal{A}^i_{\mu\nu}$ is defined via commutator of the covariant derivative[2]

$$D^{\mathcal{A}}_\mu \equiv \partial_\mu + i g_{\mathcal{A}} A_\mu \equiv \partial_\mu + i g_{\mathcal{A}} T_i \mathcal{A}^i_\mu\,, \tag{2.2}$$

[1] As always, every field and the operators induced thereof depend on the spacetime coordinates x.

[2] The partial derivative is implicitly multiplied by an identity matrix of necessary dimensions in order to match the representation of the generators.

so that

$$A_{\mu\nu} = (ig_{\mathcal{A}})^{-1}\left[D^{\mathcal{A}}_{\mu},\, D^{\mathcal{A}}_{\nu}\right] = \partial_{[\mu}A_{\nu]} + ig_{\mathcal{A}}\left[A_{\mu},\, A_{\nu}\right] = \left(\partial_{[\mu}\mathcal{A}^{c}_{\nu]} - g_{\mathcal{A}}f^{abc}\mathcal{A}^{a}_{\mu}\mathcal{A}^{b}_{\nu}\right)T_{c}\,. \quad (2.3)$$

The gauge field $\mathcal{A}^{i}_{\mu}$ corresponds to one of the fields G^{i}_{μ}, W^{i}_{μ} or B_{μ} in (2.2) and (2.3), and the gauge field strength tensor $\mathcal{A}^{i}_{\mu\nu}$ corresponds to one of the tensors $G^{i}_{\mu\nu}$, $W^{i}_{\mu\nu}$ or $B_{\mu\nu}$ in (2.1). The quantities $g_{\mathcal{A}}$, T_a and f^{abc} refer to the (real) gauge coupling, generators (in a particular representation) and structure constants[3] of a gauge group, respectively. The structure constants f^{abc} are defined by the commutation relation $[T_a,\, T_b] = if^{abc}T_c$. The non-Abelian generators are expected to follow the Killing form $\mathrm{Tr}(T_aT_b) = \frac{1}{2}\delta_{ab}$ so that

$$\mathrm{Tr}\, A^2 = \mathrm{Tr}(A_{\mu\nu}A^{\mu\nu}) = \mathcal{A}^{i}_{\mu\nu}\mathcal{A}^{\mu\nu}_{j}\,\mathrm{Tr}(T_iT_j) = \frac{1}{2}\mathcal{A}^{i}_{\mu\nu}\mathcal{A}^{\mu\nu}_{i} \quad (2.4)$$

holds. Given the local gauge transformation law $\psi \to \mathcal{U}(\boldsymbol{\alpha})\psi$ specified by the action of a unitary operator $\mathcal{U}(\boldsymbol{\alpha}) = e^{i\alpha_i(x)T_i}$ on a field ψ, the gauge field itself is required to transform according to

$$A_{\mu} \to A'_{\mu} = \mathcal{U}(\boldsymbol{\alpha})A_{\mu}\mathcal{U}(\boldsymbol{\alpha})^{\dagger} + ig^{-1}_{\mathcal{A}}\left(\partial_{\mu}\mathcal{U}(\boldsymbol{\alpha})\right)\mathcal{U}(\boldsymbol{\alpha})^{\dagger} \;\Rightarrow \quad (2.5)$$

$$\mathcal{A}^{i}_{\mu} \to \mathcal{A}^{i'}_{\mu} = \mathcal{A}^{i}_{\mu} - g^{-1}_{\mathcal{A}}D^{\mathcal{A}}_{\mu}\alpha_i + \mathcal{O}(\alpha^2_i)\,, \quad (2.6)$$

to preserve the gauge invariance of the Lagrangian density. Because the covariant derivative is expected to transform according to $D^{\mathcal{A}}_{\mu} \to D^{\mathcal{A}'}_{\mu} = \mathcal{U}(\boldsymbol{\alpha})D^{\mathcal{A}}_{\mu}\mathcal{U}(\boldsymbol{\alpha})^{\dagger}$, the interaction terms in (2.1) remain gauge invariant. The same holds for the kinetic terms as well, since

$$\mathrm{Tr}\, A^2 \to \mathrm{Tr}\, A'^2 = \mathrm{Tr}(A'_{\mu\nu}A'^{\mu\nu}) = \mathrm{Tr}\left(\mathcal{U}(\boldsymbol{\alpha})A_{\mu\nu}\mathcal{U}(\boldsymbol{\alpha})^{\dagger}\mathcal{U}(\boldsymbol{\alpha})A^{\mu\nu}\mathcal{U}(\boldsymbol{\alpha})^{\dagger}\right) = \mathrm{Tr}\, A^2\,, \quad (2.7)$$

where we used cyclic property of the trace.

The SM Lagrangian density is invariant under the local transformations of the

$$G_{\mathrm{SM}} = SU(3)_C \times SU(2)_L \times U(1)_Y \quad (2.8)$$

gauge group, where the subindex C stands for color charge, L weak isospin on left-handed fields and Y the hypercharge[4], each corresponding to gauge fields $\left\{G^{i}_{\mu}\right\}^{8}_{i=1}$, $\left\{W^{i}_{\mu}\right\}^{3}_{i=1}$ and B_{μ}. When the full group acts on a multiplet ψ (i.e. fields rearranged into a tuple), the corresponding transformation can be expressed as

$$\psi \to \exp\left(ig_S\boldsymbol{\alpha}(x)\mathbf{t} + ig_2\boldsymbol{\omega}(x)\mathbf{T} + ig_1\beta(x)\frac{Y}{2}\right)\psi\,, \quad (2.9)$$

where $\mathbf{t}$ and $\mathbf{T}$ are the generators of $SU(3)_C$ and $SU(2)_L$ groups in some representation, while Y is the hypercharge of the field ψ. The matrices representing the generators are required to be hermitian and traceless. It should be noted that the structure constants of $SU(2)$ group coincide with the Levi-Civita symbol (ε_{ijk}); the $U(1)$ group has no structure constants since it is Abelian. Finally, the real-valued differentiable functions

[3] The structure constants are absent for Abelian groups and therefore can be replaced with zero.

[4] The subindex Y is commonly used to distinguish it from the residual electromagnetic gauge group $U(1)$ left by the SSB.

$\boldsymbol{\alpha}(x)$, $\boldsymbol{\omega}(x)$ and $\beta(x)$ are arbitrary coefficients which specify a group element and, therefore, a particular transformation.

The full covariant derivative therefore reads

$$D_\mu = \partial_\mu - ig_S \mathbf{G}_\mu(x)\mathbf{t} - ig_2 \mathbf{W}_\mu(x)\mathbf{T} - ig_1 B_\mu(x)\frac{Y}{2}\,, \tag{2.10}$$

where we have wrote explicitly the dependence of the gauge fields on spacetime coordinates. However, the experiments suggest that not all fermions transform the same way as in (2.9): leptons do not participate in strong interactions and right-handed fermions do not partake in weak interactions (that explains the conventional subscript of $SU(2)_L$ as it acts only on the left-handed fields). That is why both quarks and leptons in (2.1) are separated into right-handed (subscript R) and left-handed (L) Weyl fields depending on their chirality. Thus, experimental evidence suggest the following arrangement of $SU(2)_L$ multiplets:

$$Q_L \equiv \left(Q_L^{i\alpha}\right) \equiv \left(\begin{smallmatrix} u^{i\alpha} \\ d^{i\alpha} \end{smallmatrix}\right)_L \equiv \left(\left(\begin{smallmatrix} u^{\alpha} \\ d^{\alpha} \end{smallmatrix}\right)_L , \left(\begin{smallmatrix} c^{\alpha} \\ s^{\alpha} \end{smallmatrix}\right)_L , \left(\begin{smallmatrix} t^{\alpha} \\ b^{\alpha} \end{smallmatrix}\right)_L\right) , \tag{2.11}$$

$$d_R \equiv \left(d_R^{i\alpha}\right) \equiv (d_R^\alpha,\, s_R^\alpha,\, b_R^\alpha) , \quad u_R \equiv \left(u_R^{i\alpha}\right) \equiv (u_R^\alpha,\, c_R^\alpha,\, t_R^\alpha) ; \tag{2.12}$$

$$E_L \equiv \left(\begin{smallmatrix} \nu^i \\ e^i \end{smallmatrix}\right)_L \equiv \left(\left(\begin{smallmatrix} \nu_e \\ e \end{smallmatrix}\right)_L , \left(\begin{smallmatrix} \nu_\mu \\ \mu \end{smallmatrix}\right)_L , \left(\begin{smallmatrix} \nu_\tau \\ \tau \end{smallmatrix}\right)_L\right) , \quad E_R \equiv \left(e_R^i\right) \equiv (e_R,\, \mu_R,\, \tau_R) . \tag{2.13}$$

The first line is for left-handed up and down type quarks; the second line shows right-handed up and down type quarks; the third line denotes left- and right-handed leptons. As always, summation over generation index i and flavor is assumed in (2.1). The $SU(3)$ color indices α run over the three color charges ($\alpha \in \{\mathrm{R},\, \mathrm{G},\, \mathrm{B}\}$). The notation above assumes interaction eigenstates in contrast to mass eigenstates.

One can already anticipate that the representation of G_{SM} is different for each fermion group above; in its full glory the representation can be written as

$$\mathrm{rep}(G_{\mathrm{SM}}) = \underbrace{(\mathbf{3}, \mathbf{2}, {}^1\!/_3)}_{Q_L} \oplus \underbrace{(\mathbf{3}, \mathbf{1}, {}^4\!/_3)}_{u_R} \oplus \underbrace{(\mathbf{3}, \mathbf{1}, -{}^2\!/_3)}_{d_R} \oplus \underbrace{(\mathbf{1}, \mathbf{2}, -1)}_{E_L} \oplus \underbrace{(\mathbf{1}, \mathbf{1}, -2)}_{E_R} , \tag{2.14}$$

where the first entry of each triplet denotes $\mathrm{rep}(SU(3)_C)$, the second entry $\mathrm{rep}(SU(2)_L)$ and the third shows the hypercharge Y under $U(1)_Y$. The representation (2.14) also dictates the representation of the generators in (2.9): $\mathrm{rep}(SU(3)_C)$ is fundamental (Gell-Mann matrices[5]) on triplets, but is missing on doublets and singlets; $\mathrm{rep}(SU(2)_L)$ is adjoint on triplets, fundamental (Pauli matrices[5]) on doublets and missing on singlets; $\mathrm{rep}(U(1)_Y)$ is always an identity matrix of necessary dimensions, but the hypercharges differ across the described fermion groups.

In conclusion:

- there are eight gluon fields arising from $SU(3)_C$ and four vector bosons coming from $SU(2)_L \times U(1)_Y$ covering the EW sector;
- two weak flavors in three different $SU(3)_C$ colors per one mass generation amounts to 18 four-component Dirac spinors (or 36 left- and right-handed two-component Weyl spinors) for quarks, from which

[5] The matrices are multiplied by a factor of 1/2 so that the Killing form would be respected.

  - the left-handed quarks form nine doublets (three flavor doublets per each color) under $SU(2)_L$ and six color triplets (one per flavor) under $SU(3)_C$;
  - the right-handed quarks form only color triplets for the six flavors, but the up- and down-type quarks have different hypercharges;
- the three generations of leptons amounts to three Dirac spinors (or six left- and right-handed Weyl spinors), from which
  - the left-handed leptons combined with left-handed neutrinos form three $SU(2)_L$ doublets;
  - the right-handed leptons are singlets under the non-Abelian groups;
- we assume that there are no right-handed neutrinos.

Thus, there are 45 Weyl fields in total, plus their adjoint versions.

2.2. The Higgs boson

So far we have omitted mass terms in the SM Lagrangian density (2.1) for multiple reasons. First, gauge invariance forbids massive gauge bosons, but the measurements show that the weak-interaction gauge bosons do have masses. In other words, we cannot add the term $\frac{1}{2}m_{Wi}^2 W_\mu^i W_i^\mu$ to the Lagrangian density as it would otherwise spoil the gauge invariance (compare to (2.6)):

$$\frac{1}{2}m_{Wi}^2 W_\mu^i W_i^\mu \to \frac{1}{2}m_{Wi}^2 W_\mu^{i'} W_i^{\mu'} = \frac{1}{2}m_{Wi}^2 \left(W_\mu^i - g_2^{-1}\partial_\mu \alpha_i + \varepsilon_{ijk} W_\mu^j \alpha_k\right) \times \\ \times \left(W_i^\mu - g_2^{-1}\partial^\mu \alpha_i + \varepsilon_{ilm} W_l^\mu \alpha_m\right) \neq \frac{1}{2}m_{Wi}^2 W_\mu^i W_i^\mu \,. \tag{2.15}$$

Similar reasoning holds for the charged fermions as well: a simple inclusion of Dirac mass terms in the Lagrangian density (2.1) would break its gauge invariance, even though they are measured to be massive. To see why, consider a generic fermion mass term[6]

$$m\bar{\psi}\psi = m(\bar{\psi}_L + \bar{\psi}_R)(\psi_L + \psi_R) = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \,. \tag{2.16}$$

Since all left-handed fermions transform differently compared to their right-handed counterparts (see (2.14)), we conclude that such mass term would indeed break the gauge invariance of (2.1).

Finally, massive $SU(2)_L$ admits longitudinal WW-scattering the total cross section of which grows like s^2 (with increasing c.o.m energy $\sqrt{s}$), thereby violating perturbative unitarity at $\sim 1\,\text{TeV}$ scale by crossing the Froissart bound [25] (for a brief summary see e.g. [26, section 2.3]).

The said problems are solved by considering a complex scalar Higgs field ϕ in an $SU(2)_L$-doublet. The most general Lagrangian density that satisfies these requirements reads

$$\mathcal{L}_\phi = |D\phi|^2 - V(\phi) \quad \text{with} \quad V(\phi) = \mu^2 |\phi|^2 + \lambda \left(|\phi|^2\right)^2 \,, \; \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \,, \tag{2.17}$$

[6] Note that $\bar{\psi}_L\psi_L = \left(\bar{\psi}\frac{\mathbb{1}+\gamma^5}{2}\right)\cdot\left(\frac{\mathbb{1}-\gamma^5}{2}\psi\right) = \frac{1}{4}\bar{\psi}\underbrace{\left[\mathbb{1} - \left(\gamma^5\right)^2\right]}_{=0}\psi = 0$; the same argument holds for $\bar{\psi}_R\psi_R$.

where each of the fields $\{\phi_i(x)\}_{i=1}^4$ is real-valued. Unitarity requires the coupling λ to be real and per vacuum stability condition it also must be positive because otherwise the Hamiltonian density $\mathcal{H}_\phi \supset \lambda\left(|\phi|^2\right)^2$ would be unbounded from below and possesses no ground state. Let us consider the field values $\phi' \equiv \arg\min V(\phi)$ at which the potential is minimized:

$$\left|\frac{\mathrm{d}V(\phi)}{\mathrm{d}|\phi|^2}\right|_{\phi'} = \mu^2 + 2\lambda|\phi'|^2 = \mu^2 + \lambda\sum_{i=1}^4 \phi_i'^2 \stackrel{!}{=} 0 \;\Rightarrow\; v^2 \equiv \sum_{i=1}^4 \phi_i'^2 = -\frac{\mu^2}{\lambda} \geqslant 0\,. \tag{2.18}$$

Evidently, if $\mu^2 > 0$ the potential acquires minimum at $|\phi|^2 = 0$, which is exemplified in fig. 2.1. This is uninteresting because it would only introduce a new particle with mass μ to the SM, which does not solve any of the aforementioned problems. We have a more compelling case, though, if μ^2 is negative: the minimum of the potential proves to be non-zero, which consequently means that in the ground state, ϕ reaches a non-zero vacuum expectation value (VEV):

$$\langle\phi\rangle = \langle 0\,|\,\phi\,|\,0\rangle \neq 0\,, \tag{2.19}$$

i.e. the state of lowest energy is not at $|\phi| = 0$, but near the minimum of the potential $V(\phi)$. In order to bring out a particle with vanishing VEV, we can use the three d.o.f of $SU(2)_L$ to gauge away all but one field in ϕ and expand its quantum fluctuations $h(x)$ around the VEV of the Higgs field:

$$\phi(x) \stackrel{!}{=} \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x)\end{pmatrix}\,, \tag{2.20}$$

where $\langle h\rangle = 0$, thus setting

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v\end{pmatrix}\,, \tag{2.21}$$

as intended. This gauge choice is commonly known as unitary gauge. The field h is given a particle interpretation: it is the Higgs boson. Note that the gauge transformation (2.9) with the coefficients

$$\omega_1(x) = \omega_2(x) = 0 \quad \text{and} \quad g_2\omega_3(x) = g_1\beta(x) \tag{2.22}$$

on the field ϕ would still leave the VEV (2.21) intact. These d.o.f correspond to a massless photon as we will later show. The true vacuum state broke the overall $SU(2)_L \times U(1)_Y$ symmetry, although the Lagrangian density (2.1) still remains invariant under $SU(2)_L \times U(1)_Y$ transformations. Therefore, we refer to this phenomenon as spontaneous symmetry breaking (SSB).

Let us expand the kinetic term of $\mathcal{L}_\phi$ (c.f. appendix A.2 for more detailed derivation):

$$|D\phi|^2 = \frac{1}{2}(\partial h)^2 + \frac{g_2^2}{4}W^-W^+(v+h)^2 + \frac{g_1^2+g_2^2}{8}Z^2(v+h)^2\,, \tag{2.23}$$

where we have set $Y = 1$ in order to keep the VEV and the field h electrically neutral (see the reasoning in appendix A.1). The expansion (2.23) used the following identifications for the $W^\pm$ and Z boson:

$$W^\pm \equiv \frac{W_1 \mp iW_2}{\sqrt{2}}\,, \quad Z \equiv \frac{g_2W_3 - g_1B}{\sqrt{g_1^2+g_2^2}}\,, \quad A \equiv \frac{g_1W_3 + g_2B}{\sqrt{g_1^2+g_2^2}}\,. \tag{2.24}$$

The fourth field A, orthogonal to Z, does not couple to h nor does it obtain mass from the Higgs field, hence we interpret it as the photon field. The quadratic terms of (2.23) result in the masses of $W^{\pm}$ and Z:

$$m_W = v\frac{g_2}{2} \quad \text{and} \quad m_Z = v\frac{\sqrt{g_1^2+g_2^2}}{2}\,, \tag{2.25}$$

or equivalently the couplings $g_2 = 2m_W/v$ and $g_1 = 2\sqrt{m_Z^2 - m_W^2}/v$, thereby implying $2m_W > m_Z > m_W$ if $g_1 < g_2$. The kinetic part also incorporated additional interactions with the Higgs particle h. Although the WW-scattering with the new fields (2.24) still possess divergence, it can be canceled by the additional contribution involving the exchange of the Higgs particle. The last consequence of (2.17) comes from the potential term $V(\phi)$ which, by recognizing $\mu^2 = -\lambda v^2$, yields quartic and cubic self-interactions, as well as the mass term (c.f. appendix A.3):

$$\mathcal{L}_{V\phi} = -\frac{\lambda}{4}h^4 - \lambda v h^3 - \lambda v^2 h^2\,, \tag{2.26}$$

where we have omitted constants due to VEV. The Higgs mass simply reads

$$m_h = v\sqrt{2\lambda}\,. \tag{2.27}$$

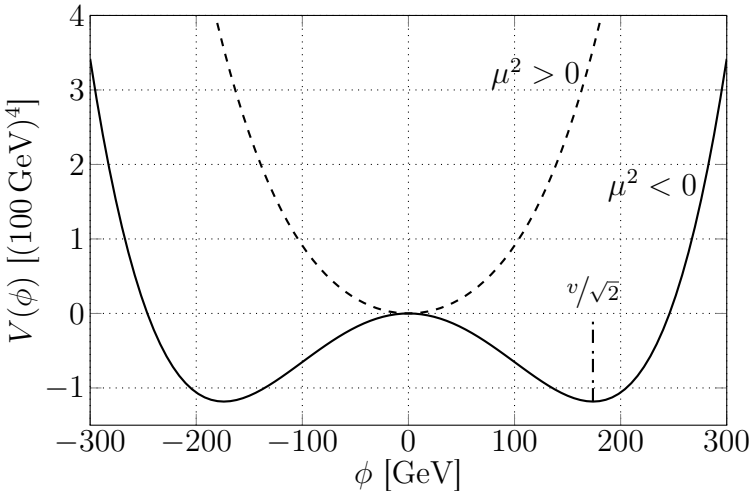


Figure 2.1: Higgs potential $V(\phi)$ for the cases $\mu^2 > 0$ (dashed) and $\mu^2 < 0$ (solid). The parameters $|\mu| \approx 88.39\,\text{GeV}$ and $\lambda \approx 0.129$ are calculated from the minimum requirements $\mu^2 = -\lambda v^2$ and (2.27), obtained from the measured values $m_h \simeq 125.0\,\text{GeV}$ and $v \simeq 246\,\text{GeV}$ [24]. The minimum of $V(\phi)$ lies at $v/\sqrt{2}$, in accordance with (2.21).

The Lagrangian density (2.1) admits all kinds of renormalizable terms which remain invariant under the gauge transformation (2.9). This allows us to recover lepton masses if we add Yukawa couplings involving Higgs and leptons:

$$\mathcal{L}_{\ell,\text{Yukawa}} = -y_{\ell i}\bar{E}_R\phi E_L + \text{h.c.}\,, \tag{2.28}$$

where $y_{\ell i}$ refers to (real) Yukawa coupling constant the mass dimension of which is zero. Once SSB has taken place, we may replace the Higgs doublet ϕ with (2.20), yielding

$$\mathcal{L}_{\ell,\text{Yukawa}} = -\frac{y_{\ell i}}{\sqrt{2}}\left[(v+h)\bar{e}^i_R e^i_L + \text{h.c.}\right] \overset{(2.16)}{=} -\frac{y_{\ell i}}{\sqrt{2}}(v+h)\bar{e}^i e^i\,, \tag{2.29}$$

from which the masses of the leptons can be easily read off:

$$m_{\ell i} = v\frac{y_{\ell i}}{\sqrt{2}}\,. \tag{2.30}$$

The mechanism also predicts new interactions between the Higgs and leptons with the coupling strength $m_{\ell i}/v$. The mass and interaction eigenstates of the leptons coincide.

Finally, the quark masses are obtained by introducing another gauge-invariant term which models Higgs coupling to quarks:

$$\mathcal{L}_{q,\text{Yukawa}} = -y^u_{ij}\bar{u}^i_R\bar{\phi}^\dagger Q^j_L - y^d_{ij}\bar{d}^i_R\phi^\dagger Q^j_L + \text{h.c.}\,, \tag{2.31}$$

where $\bar{\phi}$ stands for charge-conjugated Higgs doublet (see appendix A.1 for more information). The 3×3 complex matrices y^u_{ij} and y^d_{ij} amount to 36 coupling parameters in total. Because the Higgs field has non-zero VEV, we get quark mass matrices

$$\mathcal{M}^u_{ij} = v\frac{y^u_{ij}}{\sqrt{2}} \quad \text{and} \quad \mathcal{M}^d_{ij} = v\frac{y^d_{ij}}{\sqrt{2}}\,, \tag{2.32}$$

both of which can be diagonalized by means of biunitary transformations:

$$\hat{\mathcal{M}}^u \equiv U_R^\dagger \mathcal{M}^u U_L = \text{diag}(m_u,\, m_c,\, m_t)\,, \quad \hat{\mathcal{M}}^d \equiv D_R^\dagger \mathcal{M}^d D_L = \text{diag}(m_d,\, m_s,\, m_b)\,. \tag{2.33}$$

The four matrices $\{U,\, D\}_{L,R}$ relate interaction eigenstates $\{u,\, d\}_{L,R}$ to the mass eigenstates $\{u',\, d'\}_{L,R}$ via

$$u_{L,R} = U_{L,R}u'_{L,R}\,, \quad d_{L,R} = D_{L,R}d'_{L,R}\,. \tag{2.34}$$

Thus, the Yukawa contribution (2.31) is completely diagonalizable and therefore easily expressed in the mass eigenstates of the quarks (for more convincing explanation see A.4):

$$\mathcal{L}_{q,\text{Yukawa}} = -m_{qi}\bar{q}'_i q'_i - y^i_{qqh}\bar{q}'_i q'_i h\,, \tag{2.35}$$

where $\{q_i\}_{i=1}^6$ stands for all quark flavors. The Yukawa coupling y^i_{qqh} is proportional to the masses:

$$y^i_{qqh} = \frac{\sqrt{2}m_{qi}}{v}\,, \quad \text{or equivalently} \quad m_{qi} = v\frac{y^i_{qqh}}{\sqrt{2}}\,. \tag{2.36}$$

When we apply the same change of basis to the interaction term of the left-handed quarks, we obtain

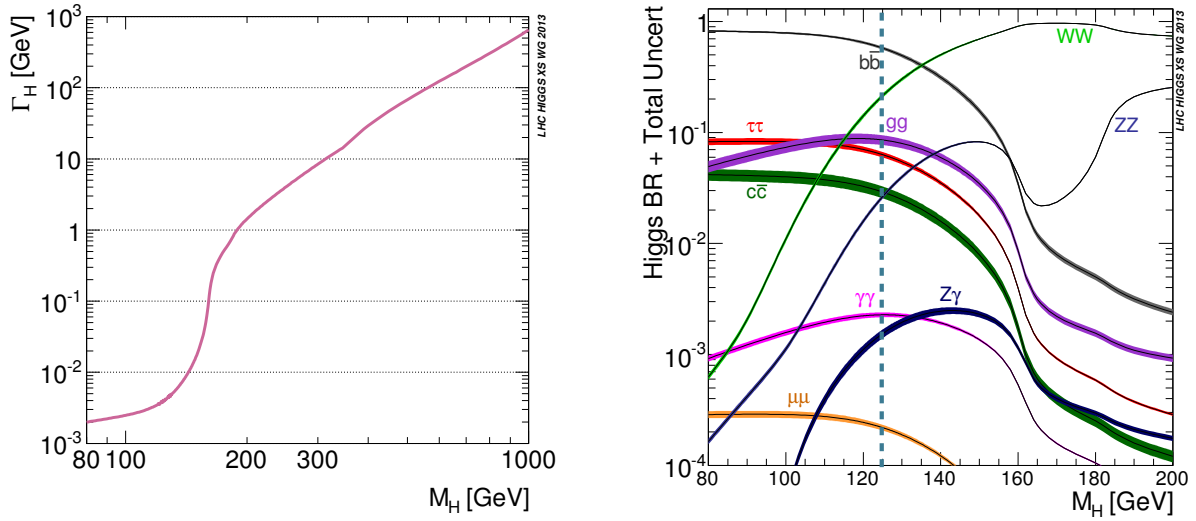
$$i\bar{Q}_L \not{D} Q_L \supset \frac{g_2}{\sqrt{2}}(\bar{u}'_L V \not{W}^+ d'_L + \bar{d}'_L V^\dagger \not{W}^- u'_L)\,. \tag{2.37}$$

The mass and interaction eigenstates are related by the Cabbibo-Kobayashi-Maskawa (CKM) matrix, which is defined by $V = U_L^\dagger D_L$. Since the CKM matrix is not equal to an identity matrix, we conclude that the physical quarks change flavor when interacting with the $W^\pm$-bosons. The interactions involving Z boson do not change flavor of the quarks or leptons, though, because the associated operator $T_3 - s_w^2 Q$ is diagonal.

In conclusion, the Higgs boson simultaneously solves multiple problems of the SM the fermions and weak gauge bosons gain masses through Higgs mechanism. The WW-scattering in this framework respects the unitary bound thanks to additional Higgs exchange which suppresses the cross section back to perturbatively unitary scales. Interestingly, VEV of the Higgs field has given masses to all massive particles (c.f. (2.25), (2.27), (2.30), (2.36)), while implying new interactions with the Higgs boson arising from the field. The final Lagrangian density of the SM includes contributions from (2.1), (2.23), (2.17), (2.29) and (2.35).

2.3. Higgs phenomenology

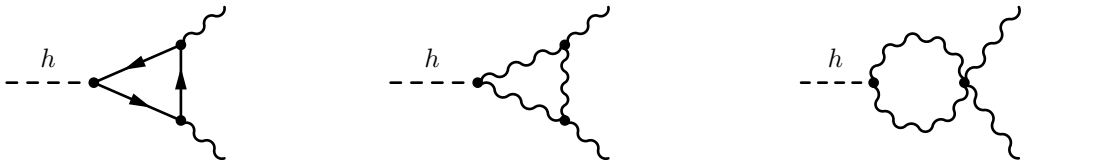
The SM we have described so far predicts various interactions involving the Higgs boson, its decay width Γ_h and branching ratios, which were commonly parametrized by its mass m_h before its discovery, as shown in fig. 2.2 (for the definitions see appendix A.5). Given that the Higgs mass has been measured to $m_h = 125.09\,\text{GeV}$, the SM predicts width of $\Gamma_h^{\text{SM}} = 4.10\,\text{MeV}$ [27, p 29]. Although instrumental resolution of $\mathcal{O}(1.5\,\text{GeV})$ limits direct detection of off-shell Higgs production, ATLAS and CMS collaborations have nevertheless studied this property by investigating its branching ratios and model-dependent off-shell couplings, thereby providing upper limits compatible with the SM estimation [24, p 193][28].



(a) (b)

Figure 2.2: Dependence of Higgs width (a) and branching ratios (b) on Higgs mass [29]. The theoretical uncertainties in (b) are indicated as bands. Explicit values are given in table D.3 in the appendix.

Higgs boson is a short-lived particle that does not appear in initial or final states at colliders. The above picture confirms the picture presented in the previous section with Higgs boson decaying into same-flavor opposite sign (SFOS) pair of fermions or into a pair of vector bosons $W^{\pm}W^{\mp}$ or ZZ. However, according to the SM the Higgs does not couple to massless bosons directly, and the gg, $\gamma\gamma$ and $Z\gamma$ pair production via Higgs channel takes place through fermion or vector boson loop as shown in fig. 2.3.

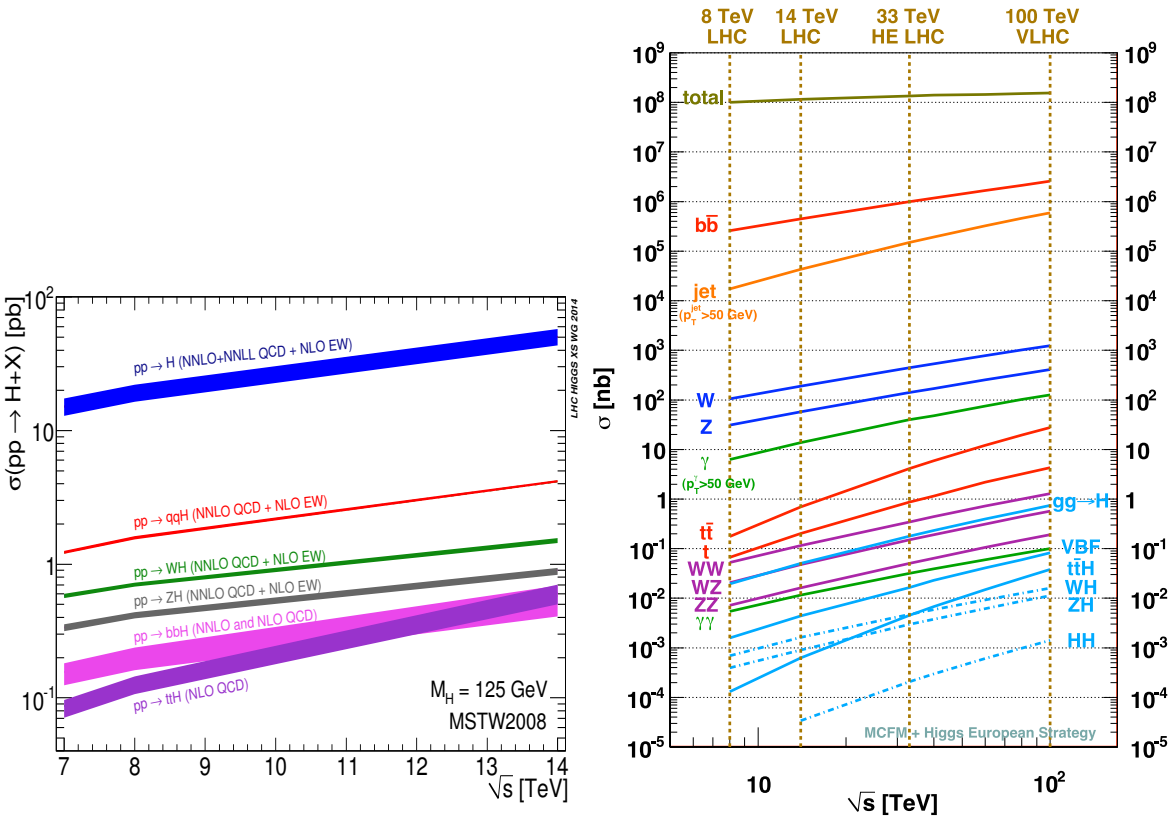


(a) Fermion loop. (b) Cubic vector boson loop. (c) Quartic vector boson loop.

Figure 2.3: Feynman diagrams for loop-induced decays $h \to gg/\gamma\gamma/Z\gamma$.

As we already know from (2.30) and (2.36), the interaction strength between Higgs and fermions is proportional to the fermion masses, so, naively, one would assume that

the Higgs would most likely decay into top-antitop pair as they have the largest mass amongst the SM fermions. However, the mass of a single top quark is $m_t = 173.21\,\text{GeV}$ ($\Gamma_t = 1.41\,\text{GeV}$) which would require either the decay products or the Higgs itself to be very off-shell, rendering the probability of such process basically zero. Thus, the dominating fermionic decay channel is actually $h \to b\bar{b}$, followed by $h \to \tau^+\tau^-$, $h \to c\bar{c}$ and $h \to \mu^+\mu^-$. The latest numerical evaluations of fermionic decays include $\mathcal{O}(\alpha_S^4)$ massless QCD corrections[7] ($b\bar{b}$ and $c\bar{c}$ channels) and EW corrections at NLO ($b\bar{b}$, $c\bar{c}$ and $\tau^+\tau^-$), leaving about 0.1% and 1% uncertainty because of truncated perturbative series [30, 31]. The $b\bar{b}$ and $\tau^+\tau^-$ decay modes are rather disfavored due to large backgrounds and poor mass resolution (10-15% of m_h).



(a) (b)

Figure 2.4: (a) dependency of different Higgs production cross sections on c.o.m energy in proton-proton collisions, with theoretical uncertainties shown as bands [32, p 276]: total inclusive (blue), VBF (red, magenta), Higgs *Strahlung* (green, grey) and $t\bar{t}h$ production (purple); (b) total cross sections of various processes due to pp collisions at the LHC, for perspective [33].

The direct decay of Higgs boson into massive vector bosons requires at least one of the decay products to be off-shell[8] because on-shell diboson mass exceeds the Higgs boson mass. Additional neutrinos in leptonic $h \to WW^*$ decays increase the mass resolution of reconstructed Higgs up to 20%. In diphotonic ($h \to \gamma\gamma$) and fully leptonic $h \to ZZ^* \to 4\ell$

[7] The coupling parameter α_S is defined by the relation $\alpha_S \equiv \frac{g_S^2}{4\pi}$, analogous to the definition of (electromagnetic) fine-structure constant $\alpha \equiv \frac{e^2}{4\pi}$.

[8] Off-shell decay products are often marked with an asterisk in literature.

decays, the reconstructed Higgs mass resolution is just at the 1-2% level, which is the reason why both channels lead to the discovery of the boson.

The fermion loop of $h \to gg$ is dominated by top and bottom quarks because they are the most massive fermions, thus having the strongest coupling to the Higgs, without the requirement of being on-shell. The QCD corrections to the gluon decay channel via top loop are known up to $\mathcal{O}(\alpha_S^5)$, leaving about a 3% uncertainty due to the renormalization scale [34]. The EW corrections have been calculated up to two loops with 1% uncertainty from missing higher order contributions [35]. The W-boson loop contributes mostly to diphotonic decay, which destructively interferes with the top loop, leading to $\sim$ 30% reduction in partial decay width of the diphotonic channel [21, p 21]. The calculation has been carried out up to NLO at 1% uncertainty [36]. Finally, the $Z\gamma$ channel is also dominated by the W loop (followed by top, bottom and τ loop), the calculations of which have been implemented only at LO, leaving an estimated 5% of residual uncertainty [37].

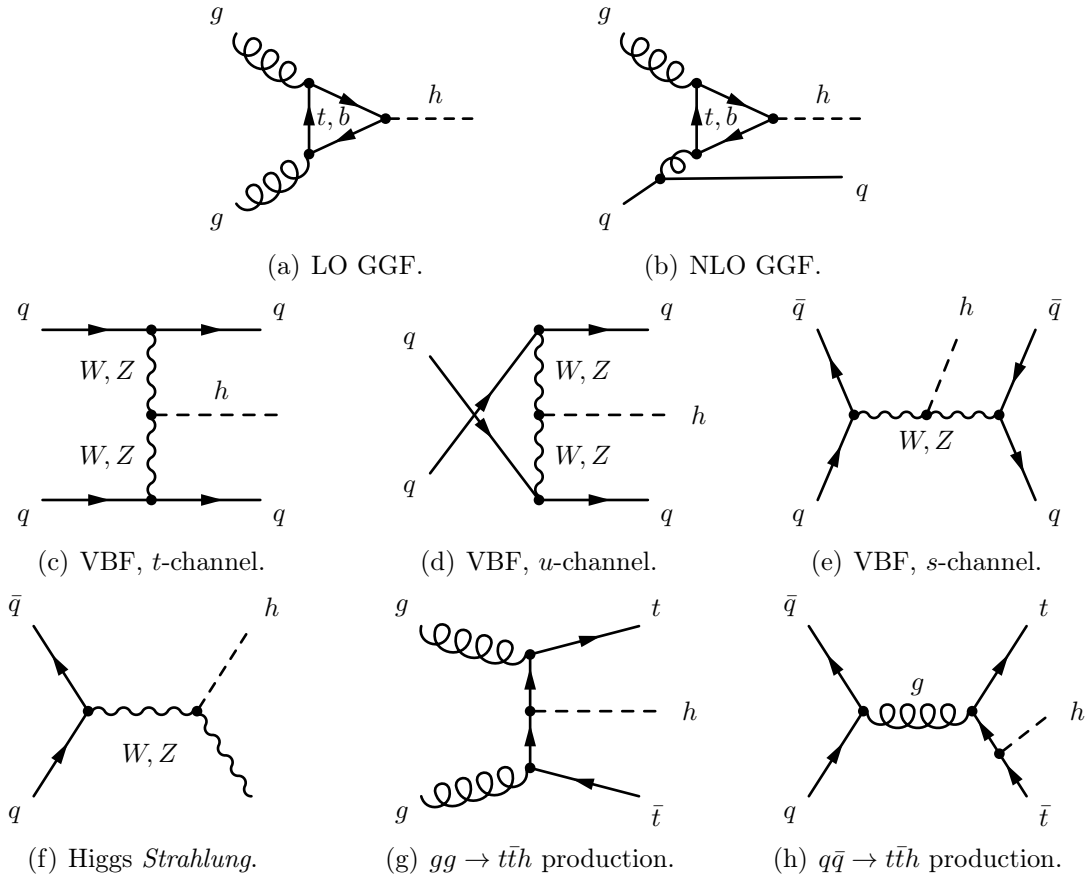


Figure 2.5: Feynman diagrams for Higgs production. Figure (b) is a prime example of initial state radiation (ISR) in which an incoming quark emits a gluon that later hadronizes in the final state, thereby producing a hadronic jet. The $b\bar{b}h$ production is similar to $t\bar{t}h$ but its cross section has dropped to that of $t\bar{t}h$ at current LHC energies (see fig. 2.2(b)). Explicit Feynman diagram corresponding to $3\ell + 1\tau_h$ signature is shown in fig. 4.1.

There are four main Higgs production mechanisms: gluon-gluon fusion (GGF) (fig. 2.5(a)), vector boson fusion (VBF) (fig. 2.5(c)-(e)), Higgs *Strahlung* or associated production with W/Z (fig. 2.5(f)) and $t\bar{t}h$ production (fig. 2.5(g)). The cross sections of each production modes are plotted as a function of c.o.m energy in fig. 2.4, exact values at current Higgs

mass and LHC energies are given in table D.1 in the appendix.

At LO, GGF comprises the same diagrams that enter the $h \to gg$ decay, which means that mostly top and bottom quarks participate in the fermion loop. This mode dominates Higgs production as it makes up about 85% of the total inclusive[9] Higgs cross section at the LHC. QCD corrections to GGF are very important because NLO contributions increase the production cross section by 80% and NNLO corrections add another 30% [24, p 176]; EW effects at NLO add mere 5% to the cross section [38]. In practice, the corrections might manifest themselves as jets in particle accelerators through ISR (see fig. 2.5(b)).

The second largest cross section at the LHC is VBF mechanism ($qq \to hqq$) in which the quarks scatter by exchanging weak gauge bosons that radiate the Higgs boson. The signature of such process typically involves two energetic hadronic jets arising from the scattered quarks, positioned back-to-back along the collision axis [24, p 177]. The cross section of VBF is about one tenth of GGF, which includes QCD (EW) contributions up to NNLO (NLO) level, reducing the residual scale uncertainties to about 2% [39, 40].

The third most probable Higgs production mechanism is Higgs *Strahlung*, or associated W/Z production with Higgs. Initial quark-antiquark pair annihilates into W or Z, which then radiates off a Higgs boson ($q\bar{q} \to hW$ and $q\bar{q} \to hZ$). If the mediated vector boson is a W, then the initiating quarks must have different flavors. Their combined production cross section is about 60% of VBF, the QCD (EW) corrections of which are known up to NNLO (NLO) [41].

The rarest main production mechanism is $t\bar{t}h$ because the initiating partons (i.e. gluons and quarks) must have a combined mass of at least $2m_t + m_h$, much higher than competing production modes. The cross section of such process makes up about one third of combined cross section of Higgs *Strahlung* at current operating energy of the LHC. Efforts at evaluating NLO QCD corrections show a moderate 20% increase in its cross section [42].

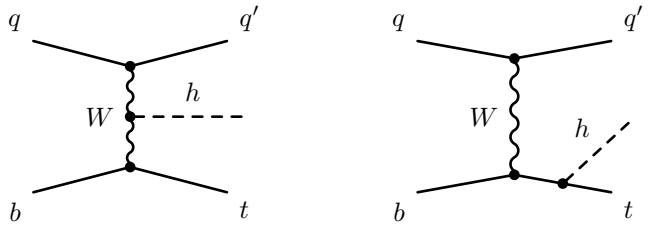


Figure 2.6: Dominant Feynman diagrams of thq process (t-channel). The bottom quark may originate from gluon splitting.

In contrast to GGF mediated by a top loop, the $t\bar{t}h$ process offers a direct access to Higgs-to-top Yukawa coupling y_t. Possible deviations from its SM value $y_t \simeq 1$ would motivate alternative explanations beyond the SM (BSM). Also, precise measurements of y_t enable the search for non-SM particles participating in loop-induced Higgs production and decay. One would argue that production of Higgs with a single top quark has lower mass threshold and therefore would be more favorable for probing y_t. However, the $pp \to thq$ production suffers destructive interference (c.f. fig. 2.6) which reduces the cross section by a factor of five w.r.t that of $t\bar{t}h$, and therefore makes it more difficult to detect than

[9] It means that all processes which involve Higgs production contribute to inclusive Higgs cross section.

$t\bar{t}h$ [32, p 741,747]. Unfortunately, $t\bar{t}h$ is insensitive to the sign of y_t – if anomalous (i.e. negative), it would hint to BSM physics. Furthermore, precision studies of y_t could also provide some insight into the stability of Higgs vacuum and BSM energy scales [43].

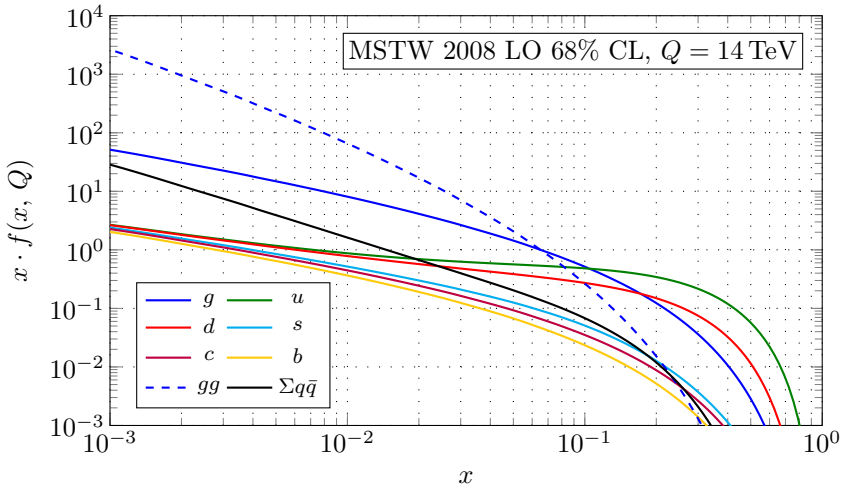


Figure 2.7: Various parton distribution functions (PDFs) evaluated at current LHC energy $\sqrt{s} = 14\,\text{TeV}$ as a function of Bjorken variable x. Recreated from [44].

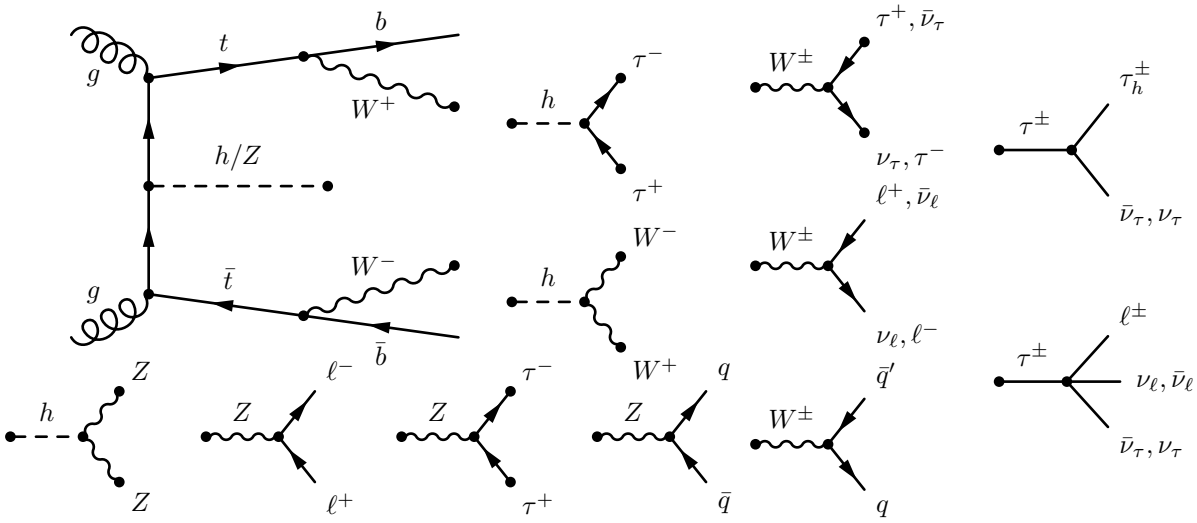


Figure 2.8: The leftmost Feynman diagram depicts the „skeleton" of $t\bar{t}h/t\bar{t}Z$ production under the assumption that $t \to bW$ (c.f. (A.21)). The small Feynman diagrams on the right and bottom show possible decay modes. Symbol ℓ stands for either electron (e) or muon (μ).

Considering the fact that LHC performs pp-collisions at $\sqrt{s} = 14\,\text{TeV}$, the most likely initial partons are gluons, followed by gq, qq' and $q\bar{q}$. There are several reasons for this: firstly, valence quarks inside a proton are u, u and d, which means that an antiquark must come from the „sea of quarks". Secondly, at as high energies as the LHC is operating today, the sea parton of an accelerating proton consist mostly of gluons. Fig. 2.7 illustrates the point: the smaller the longitudinal momentum fraction x of a colliding hadron w.r.t the c.o.m energy, the more likely it is that this fraction translates into an incoming gluon in a hard scattering event. For instance, let us consider $t\bar{t}h$ production at the LHC. This process needs a minimum energy of $2m_t + m_h$, which (albeit simplistic approximation) corresponds

to momentum ratio (also known as Bjorken variable) $x = (2m_t + m_h)/\sqrt{s} \approx 3.4 \times 10^{-2}$. According to fig. 2.7, the process (at the mass threshold) is therefore at least 100 times more likely to be initiated by a gluon pair than a quark-antiquark pair.

In this work we concentrate on the $t\bar{t}h$ production where the Higgs decays into a pair of τ leptons. The „skeleton" of this process is depicted in fig. 2.8, where we explicitly distinguish τ decays from other leptonic decays. Since the τ lepton has a mean lifetime of 2.90×10^{-13} s, it does not appear in finale states at colliders, rather it decays leptonically ($\tau^+ \to \ell^+ \nu_\ell \bar{\nu}_\tau$) in 35% of the cases or hadronically ($\tau^+ \to \tau_h^+ \bar{\nu}_\tau$) in 65% of the time (see table D.2 for precise branching ratios). We define hadronic and leptonic τ decays as mutually exclusive – any non-leptonic decay is by definition hadronic decay.

Previous studies on $t\bar{t}H$ production are summarized in table 2.1. Regarding $t\bar{t}h(\to \tau^+\tau^-)$ case involving hadronic τ decay products, only the following signal regions have previously been covered: two same-charged leptons (electrons or muons) with one hadronic τ decay product (symbolically $2\ell\text{SS} + 1\tau_h$, where SS stands for same-sign) done by the ATLAS (at $\sqrt{s} = 8, 13\,\text{TeV}$) and CMS ($\sqrt{s} = 13\,\text{TeV}$) collaborations [45–49]; and one lepton with two hadronic τ decay products (symbolically $1\ell + 2\tau_h$) by both CMS and ATLAS ($\sqrt{s} = 7, 8\,\text{TeV}$) [45, 46, 50]. Latest analysis in this production-decay mode at CMS includes new $3\ell + 1\tau_h$ category, which this thesis now focuses on [5]. Considerable background processes which have to be accounted for in this signal region (SR) include $t\bar{t}Z$ (also shown in fig. 2.8) and so-called self-background $h \to W^*W$.

	$\gamma\gamma$	$W^*W/ZZ^*/\tau^+\tau^-$	$b\bar{b}$
1ℓ	A@7, 8[45, 51]&@13[47, 52], C@7, 8[50]&@13[53, 54]	✗	A@8[45, 55]&@13[47, 56] C@7, 8[50, 57]&@13[58]
2ℓ	A@7, 8[45, 51],C@13[53]	A@8[45, 46]&@13[47, 48] C@7, 8[50]&@13[59, 60]	A@8[45, 55]&@13[47, 56] C@7, 8[50, 57]&@13[54, 58]
$2\ell + 1\tau_h$	✗	A@8[45, 46]&@13[47, 48] C@13[49]	✗
$1\ell + 2\tau_h$	✗	A@8[45, 46], C@7, 8[50]	✗
3ℓ	✗	A@8[45, 46]&@13[47, 48] C@7, 8[50]&@13[59, 60]	✗
4ℓ	✗	A@8[45, 46]&@13[47, 48] C@7, 8[50]&@13[60]	✗

Table 2.1: Summary of studies on $t\bar{t}h$ performed at the LHC classified by exclusive lepton signal regions. Letters A and C stand for ATLAS and CMS, respectively; number(s) followed by @ symbol indicate the energies at which the data had been collected and analyzed. In the context of $t\bar{t}h$ production the decay modes $WW^*/ZZ^*/\tau^+\tau^-$ give rise to multiple leptons which is why the analyses thereof are oftentimes referred to as multilepton analyses.

The final figure extracted from the analyses is measurement of the signal rate μ, defined as ratio of measured cross section to the SM prediction:

$$\mu = \frac{\sigma}{\sigma_{\text{SM}}}\,. \tag{2.38}$$

Since the Yukawa coupling y_t enters to $t\bar{t}h$ cross section at tree-level in quadrature ($\sigma_{t\bar{t}h} \propto y_t^2$), statistically significant measurements of μ can provide information about the true value of y_t.

3. Experimental setup

This section gives an overview of the Large Hadron Collider (LHC) [61] and one of its main detectors – the Compact Muon Solenoid (CMS) [62] experiment.

3.1. The LHC

The LHC, built by CERN (European Organization for Nuclear Research) between 1998 and 2008, is a hadron-hadron collider installed in the former LEP (Large Electron-Positron Collider) tunnel near Geneva, Switzerland. The repurposed accelerator ring is 27 km in circumference and located about 100 m below the ground level. The machine has taken data for analysis since April 2010, starting from $\sqrt{s} = 7\,\text{TeV}$, and has currently reached $\sqrt{s} = 13\,\text{TeV}$, just 1 TeV short of its design energy for the proton-proton (pp) collisions. The LHC primarily accelerates protons, although experiments with proton-lead (p-Pb) and lead-lead (Pb-Pb) collisions have also been conducted for periods of few weeks prior to each scheduled shutdown of the machine.

With 1232 superconducting dipole magnets guiding the beams through a circular path, the coils generating 8.3 T magnetic field have to be cooled down with liquid He^4 to a fixed temperature of 1.9 K to avoid quenching[1]. The accelerator accommodates two separate hadron beams that travel in opposite directions. Some 392 quadrupole magnets focus the beams at four interaction regions (IRs) so that they would cross inside the detectors.

The LHC has two general-purpose detectors – ATLAS and CMS – located at IR1 and IR5, respectively (see fig. 3.1). The remaining two detectors are used for dedicated research: ALICE (A Large Ion Collider Experiment) for studying Pb-Pb collisions and quark-gluon plasma, and LHCb (LHC beauty) specializing in b-physics and CP-violation, both residing at IR2 and IR8, correspondingly. Other LHC experiments smaller in scale re-use the same IRs of the larger detectors: LHCf (LHC forward) at IR1, MoEDAL (Monopole and Exotics Detector at the LHC) at IR8 and TOTEM (TOTal Elastic and diffractive cross section Measurement) at IR5.

The acceleration of hadrons proceeds through a series of different pre-accelerators not shown in fig. 3.1. The initial system is the linear accelerator LINAC2 (or LINAC3 for Pb) that generates 50 MeV protons, followed by the Proton Synchroton Booster (PSB) which pushes the energy to 1.4 GeV. The particles are next injected to the Proton Synchroton (PS), where they are lumped together into proton bunches and subsequently accelerated to 25 GeV. In the final pre-acceleration stage at the Super Proton Synchroton (SPS) their energy is increased to 450 GeV, upon which the particles are injected into the LHC ring at IR2 (also called Beam 1) and IR8 (Beam 2) such that the former beam would travel clockwise and the latter anti-clockwise. The final acceleration from 450 GeV to 6.5 TeV is performed within the LHC. Since the beam degrades over time mainly due to collisions at the IRs of the experiments and therefore would not provide sufficient event statistics for data analyses, the beam will be „dumped" or redirected to an absorbing block at Point 6 after $\sim 10\,\text{h}$ [61, p 6]. In order to maintain a high lifetime of the beams, the requirements for the vacuum are very stringent inside the beam line (10^{-10} mbar).

[1] A magnetic quench occurs when the superconducting magnet loses its superconductivity, in which large electric current causes a significant heating that damages the magnet.

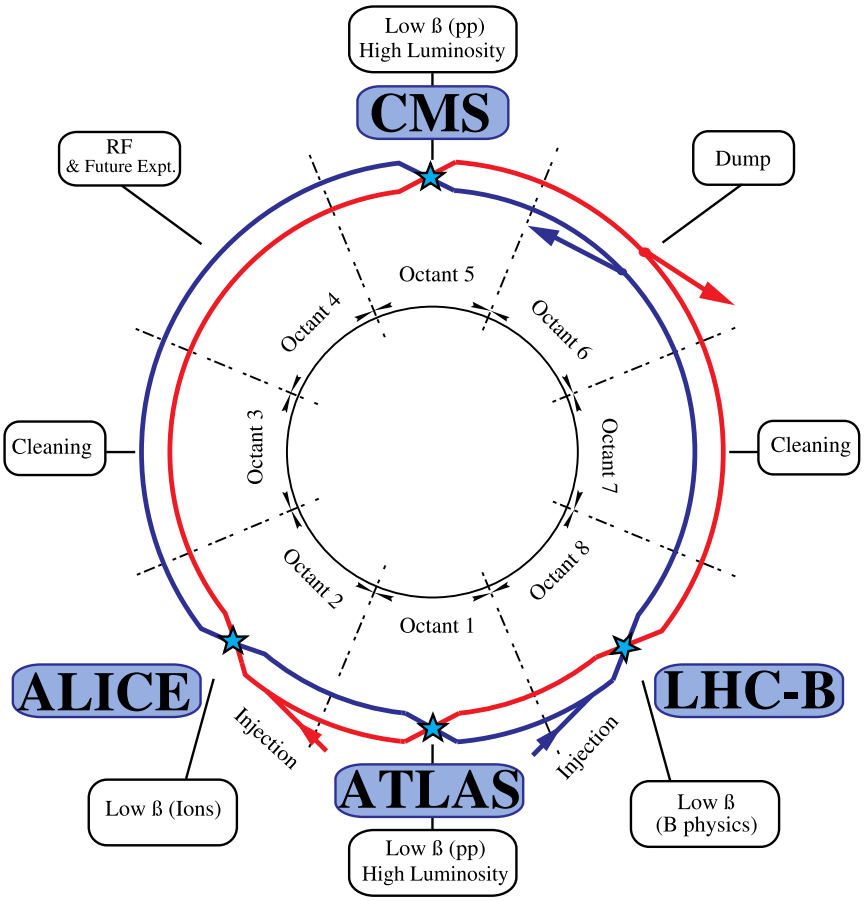


Figure 3.1: Schematic layout of the LHC, top view [61, fig. 2.1].

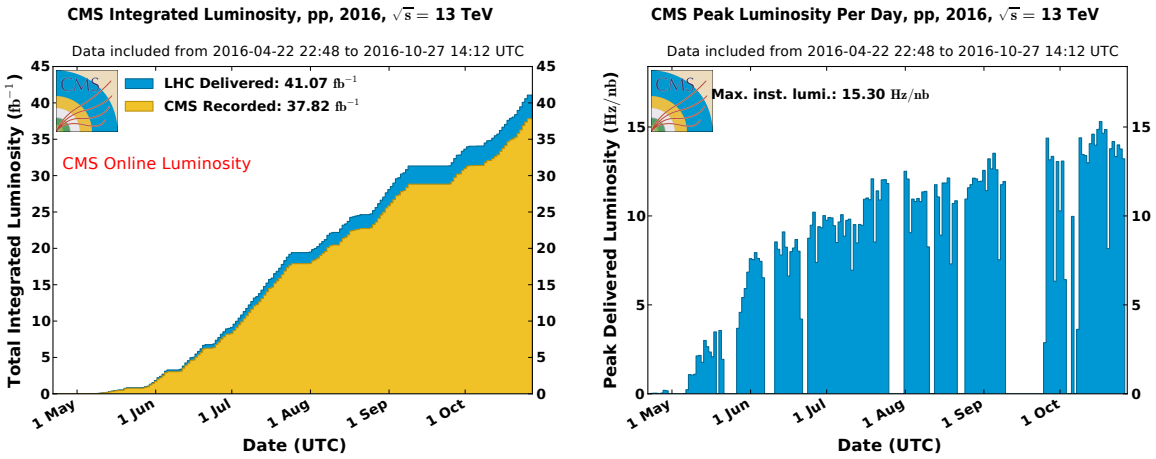
The beam consists of 2808 proton bunches separated by 25 ns per design, each bunch initially containing about 1.15×10^{11} protons. An estimate for the instantaneous collision rate is given by the formula

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L}\sigma\,, \tag{3.1}$$

where N is the number of collisions or „events", t denotes time, σ the interaction cross section and $\mathcal{L}$ the instantaneous luminosity. Simple inspection of (3.1) shows that the instantaneous luminosity has the dimensions of event flux (events per unit time and area). The total number of collisions is therefore

$$N = L\sigma \qquad \text{with} \qquad L \equiv \int \mathrm{d}t\, \mathcal{L}\,, \tag{3.2}$$

where L denotes the integrated luminosity. A peak luminosity of $1.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ was reached at the end of 2016 (see fig. 3.2(b)). The integrated luminosity serves as a measurement of collected data size and is often expressed in the units of inverse cross section (typically fb^{-1}).



(a) (b)

Figure 3.2: Luminosity results for pp collisions at 13 TeV during the 2016 data-taking period in the CMS detector [63]. Figure on the left (a) shows total integrated luminosity L, figure on the right (b) shows daily peak luminosity including maximum instantaneous luminosity of $1.53 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

Assuming Gaussian profile that the density of protons in the bunches follows in three dimensions, the instantaneous luminosity can be expressed by

$$\mathcal{L} = \frac{\gamma f N_b N_p^2}{4\pi\varepsilon_n \beta^*} \cdot \mathcal{R}(\theta_c,\, \varepsilon_n,\, \beta^*,\, s)\,, \tag{3.3}$$

where $\gamma = \sqrt{s}/(2m_p)$ is the Lorentz factor of the proton beams, m_p denotes the proton mass, f the revolution frequency of the bunches in the storage ring, N_b the number of bunches per beam, N_p the number of protons per bunch, ε_n the normalized transverse beam emittance and β^* the so-called betatron function (β-function at collision point). The reduction factor $\mathcal{R}$ captures imperfections that are due to beams crossing at non-zero angle θ_c, offset of the beam lines from their designated paths, finite length s covered by the detector along the beam line, hourglass effect (i.e. dependence of transverse beam size on the longitudinal position) and other effects due to simplifications [64]. The transverse emittance quantifies the momentum and spatial distribution of the protons inside a bunch, whereas the β-function regulates the beam envelope at the collision point. The LHC design parameters are given in table 3.1.

$\sqrt{s}$	N_b	N_p	β^*	ε_n	θ_c	f	σ_s	σ^*
14 TeV	2808	1.15×10^{11}	0.55 m	3.75 mm µrad	285 µrad	11.245 kHz	7.55 cm	16.7 µm

Table 3.1: Some LHC design parameters, valid for the ATLAS and the CMS experiments [64]. The symbols σ_s and σ^* denote the bunch length and the effective transversal size of a bunch at collision point, respectively.

According to fig. 2.4(b) and eq. (3.1), the total event rate should be 10^9 Hz. However, 40% of the pp interactions are elastic scattering events that do not actually produce any new particles [65]. Given that an estimated $N_b f \approx 3 \times 10^7$ bunch crossings happens every second, about 20 non-elastic pp interactions occur on average per bunch crossing. The

ATLAS and CMS detectors must be able to distinguish the particles originating from the $t\bar{t}h$ production event from the particles produced in other pp interactions in the same (in-time pileup, PU) or previous/subsequent (out-of-time PU) bunch crossings.

3.2. The CMS detector

The main feature of the CMS detector is a superconducting solenoid magnet of length 12.5 m and diameter 6 m that generates $|\mathbf{B}| = 4\,\mathrm{T}$ magnetic field directed along Beam 1 inside the coil. The purpose of the magnet is to bend the trajectory of a charged particle subject to the usual Lorentz force, $\mathbf{F} = (q/m) \cdot \mathbf{p} \times \mathbf{B}$, hence providing a way to measure the momentum $\mathbf{p}$ and the charge-to-mass ratio q/m of the particle. The axis of the coil coincides with the beam line. The detector components are placed around the magnet in modular manner as shown in fig. 3.3, thereby allowing to detect particles over almost the full 4π solid angle. Starting from the core and moving radially outside, the main components of the detector are tracker, calorimeters, solenoid magnet and muon detectors, all of which are enclosed within the return yoke of the magnet, a compositional steel structure weighing 12 000 t. The coordinate system is explained in appendix B.

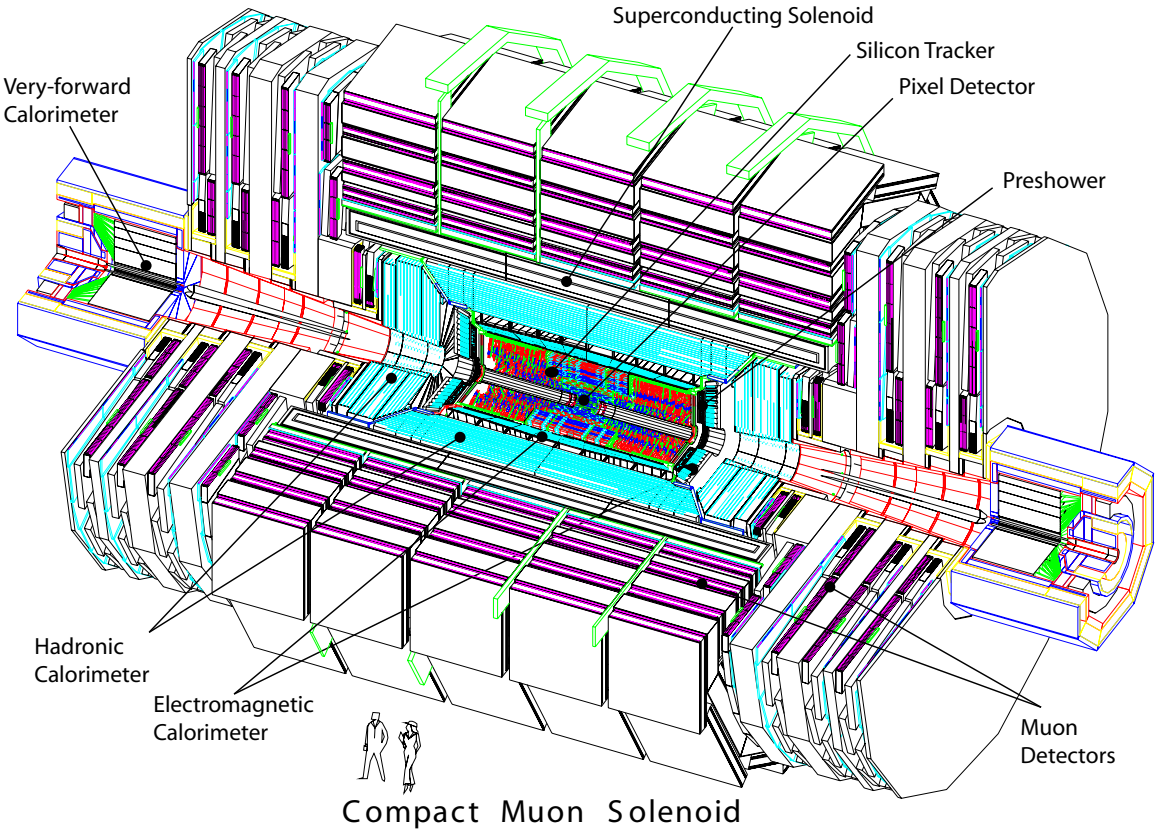


Figure 3.3: A cutaway view of the CMS detector [66]. Detector elements that form cylinders around the beam line are called the „barrel" region, whereas detector elements that are arranged in planes perpendicular to the beam line are called the „endcap" regions.

3.2.1. Silicon tracker

The silicon tracker consists of 2 to 3 layers of pixel detectors, surrounded by 10 to 12 layers of strip detectors. The pixel detector[2] consists of three 54 cm long and 285 µm thick barrel layers (BPIX), installed at radii of 4.4 cm, 7.3 cm and 10.2 cm parallel to the beam line, which are complemented by two endcap disks (FPIX) 6-15 cm in radius, installed at longitudinal distance of ±34.5 cm and ±56.5 cm from the interaction point. The first BPIX layer covers pseudorapidity region $|\eta| < 2.5$. If a charged particle passes through a pixel, it loses some energy by ionizing the Si atoms inside the pixel, upon which the created electrons and holes, subject to a bias voltage, will drift towards the electrodes, where the charge is read out. With a size of 100 µm × 150 µm, the pixels provide a very high spatial resolution that is used to „seed" track reconstruction algorithms.

The silicon strip detector is subdivided into 4-layer TIB (Tracker Inner Barrel) and 6-disk (3 per side) TID (Tracker Inner Disk) regions, both of which are surrounded by 6-layer Tracker Outer Barrel (TOB) extending to 18-disk (9 per side) TEC± (Tracker EndCap) regions. The strip detector is installed at a radial distance 20 cm to 116 cm to the beam line. In contrast to TIB and TOB regions, where the layers are parallel to the beam line, disks in TID and TEC± are positioned perpendicular to the beam line in slightly tilted (stereo) configuration and manufactured in different shapes and sizes for maximum spatial coverage and efficiency. In principle, the silicon strip detector operates similarly to the pixel detector: sensors register charge impulses generated by a charged particle passing through the silicon strip. Reduced particle flux outside of the pixel detector allows to use longer silicon micro-strips than in the pixel detector (typically 10 cm × 80 µm in TIB/TID and 25 cm × 180 µm in TOB). This design choice is also motivated by the fact that the instrumentation has to cover a larger volume which, besides costing a lot of money, would otherwise over-occupy the regions with cabling and electronics if the sensor dimensions are kept the same. The pixel and strip detectors are cooled to an operating temperature of about −10 °C, which helps to reduce damage caused by radiation.

3.2.2. Calorimeters

The purpose of calorimeters is to measure the energy of particles by completely absorbing them: photons and electrons are stopped within the volume of the electromagnetic calorimeter (ECAL), while hadrons typically traverse the ECAL and get stopped in the hadronic calorimeter (HCAL), which is located farther away from the interaction point than the ECAL. When an electron enters the ECAL, it causes a *bremsstrahlung* photon to be emitted. The *bremsstrahlung* photon in turn interacts with the material and creates an electron-positron pair. The sequence of these interactions produces a shower-like cascade of electron-positron pairs and *bremsstrahlung* photons until all of the energy has been exhausted. The energy of the initial electron decreases as $E(X) = E_0 \cdot e^{-X/X_0}$, when traversing a distance X in the ECAL. The radiation length X_0 is a material property. Good ECAL materials have large X_0, which allows for compact calorimeters. Another important material property is Molière radius, which is a measure of the transversal span

[2] At the time of writing the pixel detector has been upgraded to 4 BPIX and 3 FPIX layers (among other things like the replacement of readout chips (ROCs)) [67]. These changes should mitigate readout losses at current luminosities [68].

of the showers (w.r.t the direction of the particle initiating the shower) containing on average 90% of the energy.

In principle, HCALs work similarly: incident neutral or charged hadrons enter the material which creates a cascade of other hadrons (mostly pions) until their energy falls below the threshold for the inelastic collisions to continue. The hadrons produced in the cascade have lower energy, so that the cascade eventually stops. Analogously to radiation length in ECALs, the longitudinal development of hadronic showers in HCAL material is characterized by nuclear interaction length λ_I.

The ECAL of the CMS detector consists of 61 200 (7324) lead tungstate crystals $PbWO_4$ in the barrel (one endcap) region. The scintillating crystals are very dense ($8.28\,\mathrm{g\,cm^{-3}}$), optically transparent and resilient to radiation. Besides, they also provide a very short radiation length ($X_0 = 0.89\,\mathrm{cm}$) and small Molière radius (2.2 cm). The taper-shaped crystals in barrel (endcap) region have front-facing dimensions of 22 mm × 22 mm (28.6 mm × 28.6 mm), rear-facing dimensions of 26 mm × 26 mm (30 mm × 30 mm) and are 23 cm (22 cm) long, corresponding to $25.8X_0$ ($24.7X_0$). A single crystal covers a solid angle of $\Delta\eta \times \Delta\phi \simeq 0.0174 \times 0.0174$. The angular coverage of the barrel (endcap) detector is $|\eta| < 1.479$ ($1.479 < |\eta| < 3.0$). The ECAL also absorbs the energy of hadrons, since its interaction length amounts to $1.1\lambda_I$.

The scintillation of crystals produces light peaking in the greenish-UV region (420-430 nm), 80% of which is emitted within the bunch spacing time of 25 ns. The light is collected with two types of photodetectors that are glued to the back of the crystal: two avalanche photodiodes (APDs) with an active area of 5 mm × 5 mm in EB (ECAL Barrel) and one vacuum phototriode (VPT) with an active area of $280\,\mathrm{mm^2}$ in EE (ECAL Endcap). These photodetectors convert the collected light into electrical signals, which are further amplified, digitized and sent to off-detector electronics.

About 94% (97%) of the energy of incident electrons or photons is deposited in an array of 3×3 (5×5) crystals. Due to the silicon tracking detector that precedes the ECAL, the electrons might emit *bremsstrahlung* photons or the photons might convert into electron-positron pairs before reaching the ECAL. The 4 T magnetic field then causes the energy deposits in the ECAL to spread out in ϕ direction. The energy of the incident electron or photon is reconstructed by summing the energies deposited in adjacent crystals and corrected according to calibration data, shower position, lateral leakage (EB only), energy scale, transparency measurements of the crystals by the laser monitoring system etc [69]. In general, the resolution is around 1% of the measured energy and degrades at lower energy scales, in higher $|\eta|$ regions and in the vicinity of EB module boundaries (three per EB quadrant).

The HCAL surrounds the ECAL. It is subdivided into barrel region (HB), endcaps (HE) and forward region (HF). Unlike the ECAL which is a homogeneous calorimeter, the HCAL is a sampling calorimeter, because it is constructed from alternating layers of an absorber, which initiates particle showers, and an active medium, which produces detectable signals (i.e. the active medium „samples" the energy of particles in the showers).

The HB is segmented into 64 sectors in ϕ and 16 in η, called towers, which each cover an area of $(\Delta\phi \times \Delta\eta) = 0.087 \times 0.087$. When viewed in radial direction, each tower contains 14 layers of ~ 5 cm-thick absorbing brass separated by plastic scintillator trays. When hadrons enter the calorimeter, the scintillators emit UV light which is shifted into the green spectrum by fiber-optic waveguides of 1 mm in diameter. At the end of the

waveguides hybrid photodiodes (HPDs) are installed, which convert the light to electronic signals via the photoelectric effect. The electric signals are then amplified, digitized and transmitted to the data acquisition (DAQ) system. The combined energy resolution of the HCAL is typically $\Delta E/E \simeq 100\%/\sqrt{E} \oplus 5\%$ (in GeV).

The HF calorimeter extends the angular coverage of the detector up to $|\eta| = 5.0$. It is designed to withstand the very high radiation doses that occur close to the beam line. Instead of brass-scintillator combination that is used in other parts of HCAL, the HF uses steel as an absorber with quartz fiber optics inserted as active material. The cylindrical steel absorbant has an outer radius of 130 cm with a 25 cm diameter hole for the beam line, and is split into 18 equal sectors in the y-x plane (36 if both sides included) which are further subdivided into towers. The quartz fibers are inserted to the steel structure parallel to the beam line and measure the Cherenkov light emitted by the charged particles produced within electromagnetic or hadronic showers. The detector is housed in a hermetically sealed radiation shield made of steel and concrete, that protects the photomultipliers (PMTs) that register the light signals received from the quartz fibers from the high particle flux.

3.2.3. Muon detector

Even though a muon loses some of its energy when traversing the tracker and calorimeters, it very likely will not be stopped within these detector elements. In order to identify the muons and accurately measure their momentum, the CMS detector includes four layers of muon detectors, called stations, which are embedded in the iron return yoke of the magnet. The muon detector is separated into barrel (MB) and endcap (ME) regions, covering pseudorapidity ranges of $|\eta| < 1.2$ and $0.9 < |\eta| < 2.4$, respectively. The muon system employs tree different types of gaseous detectors: MB uses drift tube (DT) layers positioned parallel to the beam line, while ME has cathode strip chamber (CSC) layers perpendicular to the beam line. Resistive plate chambers (RPCs) are installed in the barrel and the endcap region, and provide fast signals for triggering purposes.

The muon stations contain two superlayers (SLs) that measure the muon coordinates in the r-ϕ plane (SL1 and SL3), complemented by one superlayer (SL2) that provides measurements in the r-z plane, except for the outermost muon station, which is missing this layer. SL1 and SL3 are installed at maximal radial separation to provide optimal resolution to measure the transverse momentum of a muon track. Each SL contains four layers of rectangular 13 mm × 42 mm cells, which are staggered by half a cell.

If a muon passes through the DT, it ionizes gas atoms inside the DT. The gas is a mixture of 85% Ar and 15% CO_2, and is kept at atmospheric pressure. The electrons follow the electric field, which is shaped by 5 electrodes: one anode wire (+3600 V) mounted in the middle of the cell, two anode strips (+1800 V) on the outermost sides of the cell and two cathode strips (−1200 V) on the innermost sides of the cell. The electrons drift towards the anode wire at a velocity of about 55 µm/ns, which corresponds to maximum drift time of 380 ns. The muon chambers allow to determine muon position in the r-ϕ plane with a resolution of typically 100 µm.

The operating principle of the CSCs installed in the MEs is similar to the DTs. Per muon station, six layers of anode wires and seven layers of cathode strips, both made of copper, are positioned almost perpendicularly to each other in an interleaved fashion

within a gas volume (40% Ar + 50% CO_2 + 10% CF_4). The strips run radially outwards, thus covering a constant $\Delta\phi$ section, whereas the wires run along ϕ, and define the radial coordinate. When a muon passes through these chambers and ionizes the gas, the electrons drift towards the anode, while the ions move towards the cathode strip. The muon track is reconstructed based on the position of the induced surface charge distribution on the cathode strips. Compared to the DTs, the CSCs are more capable to withstand the high radiation in the endcaps (especially in ME1). They also have faster response time and finer segmentation than DTs (about 75 µm in r-ϕ for ME1 and double that for other stations).

The key elements of the RPC are two 2 mm thick high resistive bakelite plates which are positioned parallel to each other and separated by 2 mm. A high voltage of about 9 kV is applied to the plates. The gap between the anode-cathode plates contains an ionizable gas (mostly $C_2H_2F_4$). Two such single-gap pairs with metallic readout strips between them form a double-gap RPC unit. A muon passing through such system ionizes the gas molecules in the gaps. The electrons then move towards the anode plates where the charges are read out. In the barrel region, the RPCs are mounted on either side of the DT chambers in the two innermost muon stations, and precede the DT chambers in the two outermost muon stations. In the endcap, the RPCs are installed in the three innermost muon stations in the ME, thereby covering a pseudorapidity region up to $|\eta| = 1.6$. The RPC system provides a fine timing resolution of $\lesssim 1$ ns, but fairly coarse spatial resolution of 1 cm, ideally suited for triggering purposes [70].

3.2.4. Triggers

An important component of the CMS detector is its trigger system. The purpose of the trigger system is to reduce the event rate from the 40 MHz bunch crossing frequency to an output rate of $\mathcal{O}(100\,\text{Hz})$ that is suitable for event reconstruction and data storage. Event selection at the trigger level is referred to as online selection, whereas the analysis performed on the recorded data is referred to as offline selection.

The trigger system executes the online selection in two stages: first the Level-1 (L1) trigger performs a preliminary event selection based on trigger primitives that are generated by fast electronics installed on the detector before the full detector data is read out by the DAQ and processed by the high-level trigger (HLT), which performs the final decision whether or not to keep the event. The L1 trigger reduces the event rate to 100 kHz that needs to be handled by the DAQ and HLT systems. It delivers its decision within 3.2 µs, during which time the detector data is buffered in the pipelines of the detector. The HLT is implemented using a simplified version of the offline CMS event reconstruction software [71], and is executed on a cluster of commodity computers. The trigger system is depicted in fig. 3.4.

The L1 trigger reaches its final decision in multiple steps. In the first step, a set of trigger primitives are generated based on the energy deposits in the ECAL and HCAL that exceed programmable thresholds [72]. This data is combined and transmitted to the regional trigger, which builds various types of object candidates such as electrons, photons and jets. This information is propagated to the global calorimeter trigger which calculates additional quantities such as scalar and vectorial sums of transverse energies (H_T, $\not{E}_T$), which are subsequently passed to the global trigger. In regard to muon measurements,

all muon detectors (DTs, CSCs and RPCs in both barrel and endcap) provide trigger primitives. The global muon trigger receives the information and builds muon candidates. It then sends this information to the global trigger. The global trigger makes the decision whether to disregard the event or initiate the readout of the full event information and pass the event on to the HLT. The full event information includes information from the calorimeters and muon systems with higher spatial and energy/momentum resolution, plus the data from the silicon tracking detectors, which was not used by the L1 trigger. The schematics of the L1 trigger system is depicted in fig. 3.5.

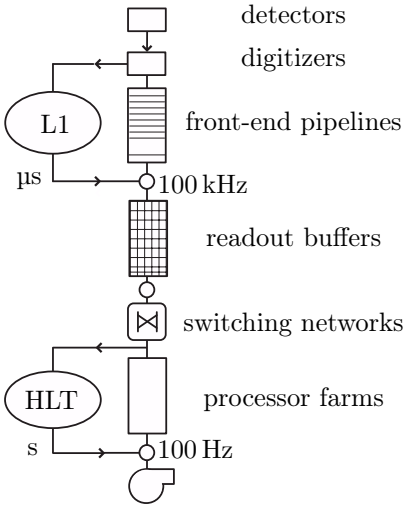


Figure 3.4: Data flow diagram of the CMS trigger and DAQ system (taken from [73]).

Once an event has been accepted by L1, the HLT algorithms perform complete reconstruction of the physics objects using a simplified version of the CMS event reconstruction software that is optimized for speed, and filter the events. There are more than 200 different triggers in the HLT, each addressing different physical aspect of the event, e.g. events containing at least one, two or three leptons of certain flavors above some p_T thresholds. These so-called trigger paths have binary outputs: either the event passed the conditions and „fired" a given trigger path, or the event failed this particular path. The event is written to permanent data storage, i.e. kept for offline data analyses if at least one trigger path is fired.

The data is stored in `RAW` and `RECO` event formats, containing the digitized hits and reconstructed objects (tracks, vertices, jets, electrons, muons etc), respectively. The information stored in the `RECO` format is further reduced for data analysis: first to the `AOD` and then to the `MiniAOD` format. The analysis presented in this thesis is based on the most compressed form, the `MiniAOD` format. The data is stored and analyzed with the ROOT framework [74].

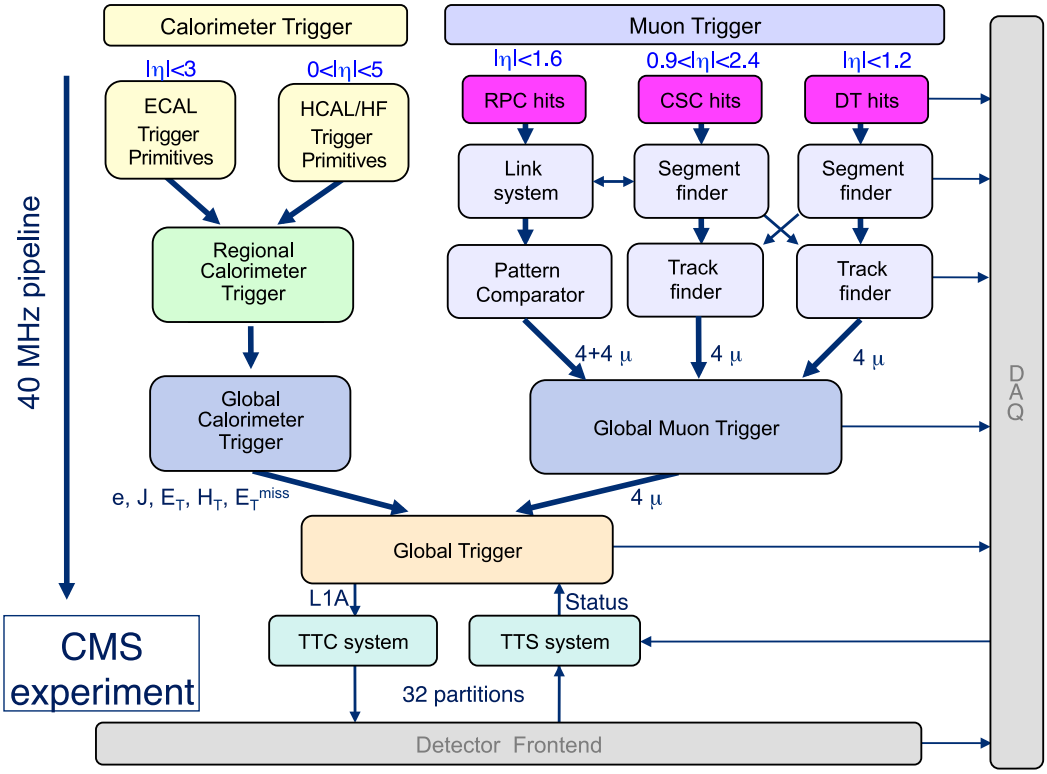


Figure 3.5: L1 trigger (adapted from [72, figure 2]).

4. The analysis

This chapter presents the analysis of $t\bar{t}h$ production in events containing three electrons or muons plus one hadronically decaying τ, referred to as $3\ell + 1\tau_h$ in final state. The strategy goes as follows: first, event selection is performed on simulated $t\bar{t}h$ signal and background events, as well as on the data. The event selection manipulates with particle objects, such as electrons, muons, jets and hadronic τ decay products. We follow the instructions published by various groups of experts in CMS who have developed algorithms for reconstructing and identifying the said objects. The events that are simulated by the Monte Carlo (MC) methods are corrected in order to account for the differences between the simulation and the data. Special care is taken to accurately model background contribution to the signal region (SR). This is followed by signal extraction with the help of boosted decision trees (BDTs), which are supervised machine learning methods designed to separate signal from background [75, 76]. Signal rates are obtained by means of a maximum likelihood fit to the data. The analysis code is made available at [3].

4.1. Data and MC samples

The data used in the analysis presented in this thesis has been recorded in 2016 by the CMS detector at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ and corresponds to an integrated luminosity of $35.9\,\text{fb}^{-1}$. The events have been collected by single and double electron (muon) triggers and by triggers that are based on an electron plus muon pair in the event. The analysis considers only the data that has been taken with all detector systems being fully operational. The corresponding „good luminosity mask" is provided by the CMS Data Quality Monitoring (DQM) and Data Certification groups and is applied at the very first step of the data analysis [77, 78]. The relevant datasets are listed in table D.4 and are available via the CMS Data Aggregation System (DAS) [79]. The events have been reconstructed with version `CMSSW_8_0_x` of the CMS reconstruction software and are stored in `MiniAOD` format [71].

In order to reliably compare measured data to SM predictions, we need accurate simulations of $t\bar{t}h$ signal and of background interactions that occur in the CMS detector. Apart from the signal process ($t\bar{t}h$), the background processes considered in this analysis include: $t\bar{t}W$, $t\bar{t}Z$, $Z+\text{jets}$, $W+\text{jets}$, $t\bar{t}(+\text{jets})$, $W+\gamma+\text{jets}$, single top, diboson (WW, WZ, ZZ) and triboson (WWW, WWZ, WZZ, ZZZ) production and rare processes such as $tt\bar{t}\bar{t}$. The signal process assumes Higgs mass of $m_h = 125\,\text{GeV}$. The events are generated from hard scattering matrix elements with MC techniques using MadGraph5_aMC@NLO [80], POWHEGv2 [81] and PYTHIA8 [82] software packages, the latter of which also models parton showers, hadronization and τ decays. All but $W+\gamma+\text{jets}$ and $t\bar{t}+\text{jets}$ matrix elements are evaluated beyond LO. Additional detector effects are simulated with the Geant4 software [83–85]. Because the production of MC samples is very time-consuming, especially considering the fact that we need many events in order to arrive at statistically reliable results whatsoever, the samples are produced, validated and published by dedicated groups. The samples used in this analysis are created in 2016 production campaign, the complete list of which is given in table D.5. The MC events contain both generator (i.e. parton) level and full reconstruction information compatible with the recorded data format (`MiniAOD`).

4.2. Object reconstruction and identification

The CMS detector allows to reconstruct and identify photons, electrons, muons, charged hadrons and neutral hadrons. The photons deposit all their energy in the ECAL and are identified as such if their energy depositions in the ECAL are not linked to any track. Electrons are identified and their tracks reconstructed from the readings in the silicon tracker and ECAL. Muons, on the other hand, combine information from the tracker and muon detector outside of the solenoid. The charged hadrons are identified by linking the energy deposits in HCAL and ECAL with the tracker information. Similarly to photons, the neutral hadrons do not interact with the tracker material and are thus identified as such from calorimeter deposits that are not linked to any track. The hadrons and photons are clustered together into jets that are initiated by either hadronic τ decays, or by hadronization of quarks or gluons. The objects used in the analysis need to pass a preselection implemented in the software package [86]. Events passing the preselection are stored in an event format (`Ntuple`), the size of which is further reduced w.r.t the `MiniAOD` event format.

4.2.1. Electrons

Global electron tracks are reconstructed by the Particle Flow (PF) algorithm [87, 88], which uses tracks reconstructed by the Kalmar filter and Gaussian Sum Filter (GSF) methods [89], and combines them with information provided by calorimeters. The basic electron reconstruction and identification is developed by a dedicated group of experts (EGamma POG) in CMS who have published guidelines for selecting electrons based on certain quality criteria.

The first step in the reconstruction of electron tracks is to find „seeds". These seeds may be found in the tracker (requiring two or three hits in the silicon pixel detector) or in the ECAL (requiring that the sum of energy deposits exceeds some threshold). The tracks are extrapolated to the ECAL in the former case or to the first pixel layer in the latter case, while considering both positive and negative charge hypotheses. Full tracks are built using the Kalman filter algorithm, which iteratively searches for hits in successive layers of the tracker.

However, electrons may already shower inside the tracker since the total tracker material (including front-end electronics, cooling, cabling and support structures) amounts to a radiation length of $0.5X_0$ to $2X_0$, depending on $|\eta|$. Thus, electrons may emit *bremsstrahlung* photons and thereby lose energy inside the tracker. Other indicators for interactions of electrons within the tracker include fewer hits in the tracker layers, sudden changes (kinks) in the curvature of the electron trajectory, and energy deposits in the ECAL that are spread out in ϕ direction by the magnetic field. In order to account for possible energy losses of electrons that traverse the tracker, tracks of electrons are fitted with the GSF method, which improves momentum resolution compared to the standard Kalman filter. *Bremsstrahlung* photons are searched in the ECAL regions in direction tangent to the GSF track at each tracker layer.

In order to distinguish electrons from jets, the EGamma POG has trained a BDT, which uses tracking information and ECAL observables of the electron candidate as input. The training has been performed for electron candidates in two p_T ranges or „bins" ($p_T < 10\,\text{GeV}$ and $p_T > 10\,\text{GeV}$) and in three η bins ($|\eta| < 0.8$, $|\eta| \geqslant 0.8$ EB, EE) [90].

The BDT outputs a value between -1 and $+1$ for each electron candidate. An output value of -1 ($+1$) means that the electron candidate is likely a jet (a genuine electron). The multivariate analysis (MVA) method is not perfect in the sense that it may misclassify electron candidates: the computed scores may lead to false positives (BDT score of $+1$ is assigned to genuine jets) or false negatives (-1 is assigned to genuine electrons) or something in-between (electron candidates are given a score of 0). For this reason, we use a working point (WP) of the BDT – if the computed BDT score of an electron candidate is above (below) the predefined WP, we assume that the candidate is a genuine electron (a jet). The EGamma POG has published recommended WPs which we employed in this analysis (see table D.8). Lower value of the WP usually means higher efficiency for correct identification of genuine electrons (i.e. true positive rate), but also a higher misidentification rate for jets (i.e. false positive rate).

We also use another MVA trained by $t\bar{t}h$ multilepton group in order to distinguish „prompt" leptons from „non-prompt" leptons [49, 59]. The prompt leptons originate from leptonic $W/Z/\tau$ decays while the non-prompt leptons come from c and b quark decays. The MVA outputs a value between 0 and 1, whereby the non-prompt leptons are expected to have lower MVA scores than the prompt leptons. The training has been performed for electrons and muons separately based on the $t\bar{t}h$ signal and $t\bar{t}$ + jets background samples. The leptons used in the training have to pass so-called „loose" preselection criteria described in table D.6. The table also lists the criteria for „fakeable" and „tight" selection of leptons. The leptons that pass the fakeable criteria (also called „fakeable leptons") are used to estimate fake background due to non-prompt leptons (explained in section 4.3.2), whereas the leptons that pass the tight criteria („tight leptons") are used in the signal event selection (described in section 4.3).

Commonly to all selection categories, we only consider electrons which are detected within $|\eta| < 2.5$ (corresponds to pseudorapidity covered by the innermost BPIX layer) and require its reconstructed tracks be sufficiently close to the primary vertex (PV) (see appendix B.4 for more information). We also employ an isolation condition (explained in appendix B.3), as well as photon conversion rejection. The latter removes electrons in case the track is near another track of opposite charge, or in case the track of the electron candidate is missing hits in the innermost layer of the silicon pixel detectors.

As for the track quality requirement, the reconstructed tracks of fakeable and tight leptons must hit all layers in the silicon tracker, whereas the loose leptons can miss at most one layer. Compactness requirements on the electron energy deposited in the ECAL are used in fakeable and tight categories only if their p_T exceeds 30 GeV (see table D.10).

4.2.2. Muons

The tracks of muons are reconstructed by considering the hits in the tracker and in the muon system. Candidates for muon tracks are reconstructed in the silicon tracker and in the muon system, which are referred to as „tracker tracks" and „standalone muons", respectively [91]. In the „outside-in" approach each standalone muon is extrapolated from the muon system to the tracker detector and matching to tracker tracks is performed. If successful, a global refit of the track using the standalone muon and tracker tracks is performed, resulting in a „global muon". In the complementary „inside-out" method, all tracker tracks with $p_T > 0.5$ GeV and $p > 2.5$ GeV are extrapolated from the tracking

detector to the muon system. Upon successful matching with at least one track stub in the muon system the pair of tracker track and muon stub is used to build a „tracker muon". If both global and tracker muons share the same tracker track, they will be merged into a single candidate by the PF algorithm which associates energy deposits in the ECAL and HCAL to the muon.

The group of muon experts in CMS (Muon POG) has developed quality criteria to separate muons from jets [92]. The muon selection employed in the analysis (shown in table D.11) requires muon to pass the loose PF muon selection, which requires the muon candidate to be reconstructed by the PF algorithm and to be associated to either a global or to a tracker muon (or both). These conditions remove accidental matching of tracks in the muon subsystem and the tracker, as well as in-flight decays of hadrons, particularly charged pions and kaons (see table D.12), and punch-through (hadronic showers that do not stop in the HCAL and „leak" into the muon system).

Similarly to the electron selection described in the previous section, we distinguish between „loose", „fakeable" and „tight" muons, depending on whether the muon candidates pass either the loose, fakeable or tight selection criteria. The purpose of such categorization is the same as for the electrons: loose muons are input to the MVA that separates prompt leptons from non-prompt lepton; fakeable muons are used in the estimation of the fake background (described in section 4.3.2); tight muons are used in the signal event selection (described in section 4.3). Only the tight muons used in the event selection need to pass the medium PF muon selection criteria, which requires the muon candidate to have at least 80% of valid hits in the silicon tracker in order to get a better estimate for its momentum. Additional requirements include either good segment compatibility[1] or so-called „good global muon" properties[2].

As instantaneous luminosity increases, less hits in the silicon strip tracker were associated to reconstructed tracks due to a suboptimal configuration of the detector electronics during the first half of the 2016 data-taking period [93]. As this inefficiency negatively impacts reconstruction of medium PF muons, the Muon POG have provided a temporary fix to mitigate the issue caused by this effect.

4.2.3. Jets

The PF objects are clustered together into jets by the so-called anti-k_t algorithm [94]. The central pieces of the algorithm are particle-particle (d_{ij}) and particle-beam (d_{iB}) distance measures,

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2} \min\left\{\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2}\right\} \quad \text{and} \quad d_{iB} = \frac{1}{k_{T,i}^2}\,, \tag{4.1}$$

where $k_{T,i}$ is transverse momentum of the i-th particle and ΔR_{ij} the angular separation between the i-th and the j-th particle (as explained in appendix B.3). The constant

[1] At first order, the segment compatibility represents a number between 0 and 1 which is calculated as a weighted sum of segment hits (1 – a hit; 0 – no hit) by the extrapolated track to the muon chambers, whereby the outermost stations are given the most weight.

[2] A good global muon must have a relatively good track fit in terms of χ^2 (a statistic describing closeness of the measurements-to-fit model), small number of discontinuities in the track (which indicate in-flight decays or overlapping of the candidate track with other reconstructed tracks) and a good matching between the tracker muon and the standalone muon.

R regulates the (approximate) size of the jets that are built by the algorithm. The algorithm first loops over all particles reconstructed by the PF algorithm and pairs thereof, and determines the distances d_{ij} and d_{iB} determined by (4.1). If the minimum distance happens to be between a particle and the beam (d_{iB}), the particle is classified as jet and excluded in subsequent iterations, hence providing a safeguard for collinear parton splitting. Otherwise the particle pair will be merged into a jet candidate, the two particles are removed from the list of particles that are still to be processed by the algorithm, and the newly built jet candidate is added as „particle" to the list instead. The 4-momentum of the jet is set to the sum of the 4-momenta of the two particles. The distances d_{ij} and d_{iB} are then recomputed. This process continues until the list of unprocessed particles is empty. The algorithm is robust against soft radiation, because the measure d_{ij} prefers merging of hard-soft pair over soft-soft pairs. If the distance between two jets is between R and $2R$, their cones are clipped based on their p_T. In our analysis we use jets built with parameter $R = 0.4$.

The energy of jets needs to be calibrated depending on p_T and η of the jet to take into account for the response of the calorimeters and to remove PU effects [95]. The calibration is applied by means of jet energy corrections (JEC). Different JEC are applied to data and MC correcting for a small imperfections in the simulation of the detector response. We use `80X_mcRun2_asymptotic_2016_TrancheIV_v6` JEC for MC and `80X_dataRun2_2016SeptRepro_v7` for data, both provided by the jet energy resolution and corrections (JERC) subgroup of CMS [96].

Jets arising from the hadronization of b-quarks are identified by the CSVv2 algorithm, developed and validated by the group of experts for b-tagging and vertex fitting (BTV) POG in CMS [97]. They define loose and medium WPs corresponding to rates of 10% and 1% for light quark jets to be misidentified as b-jets, respectively. Jets used in our analysis are required to pass $p_T > 25\,\text{GeV}$ and to be within the geometric acceptance $|\eta| < 2.4$. The principles of b-tagging are briefly explained in appendix B.4.

4.2.4. Hadronic τ decays

Hadronic τ decays (τ_h) are reconstructed by the hadrons plus strips (HPS) algorithm [98, 99]. The algorithm is seeded by jets reconstructed by the anti-k_t algorithm with $R = 0.4$. The algorithm aims to reconstruct the main τ decay modes: $h^\pm$, $h^\pm + \pi^0$, $h^\pm + 2\pi^0$ and $h^\pm h^\mp h^\pm$, where $h^\pm$ stands for either a charged pion ($\pi^\pm$) or kaon ($K^\pm$). The branching ratios of these decay modes are listed in table D.12.

Because of a large branching ratio for a neutral pion (π^0) to decay into a photon pair, and the high probability for one of the photons to convert into an electron-positron pair while traversing the tracking detector, possibly followed by *bremsstrahlung* and further conversions, the algorithm reconstructs π^0 as clusters of photons and/or electrons that are within rectangular strips. The size of the strip in ϕ direction depends inversely on the p_T of its e/γ constituents, while taking into account that the bending of $e^\pm$ tracks decreases for $e^\pm$ of high p_T. The algorithm always ends in finite steps since the maximum strip size is limited ($\Delta\phi = 0.3$ and $\Delta\eta = 0.15$).

Thus, a hadronic τ is formed from charged hadrons and π^0 candidates reconstructed by the above algorithm if the reconstruction of charged hadrons and π^0 matches one of the decay modes given before. The 4-momentum of the τ_h is determined by summing the

momenta of its constituents.

Because τ_h signature may be faked by hadronic jets, an MVA has been developed by the CMS group specializing in τ_h reconstruction (Tau POG) in order to separate τ_h from the jet background [100]. The BDT has been trained on genuine τ_h and jets and takes the following inputs: decay mode, scalar p_T sum of its constituents, and observables that are sensitive to the lifetime of the τ: d_{xy} of the leading track and distance between PV and SV (secondary vertex) formed from three charged tracks in case of 3-prong decay mode. The Tau POG has provided various WPs corresponding to different signal efficiencies.

For the purpose of this analysis, we distinguish fakeable and tight τ_h based on the WP that the τ_h passes (see table D.13). The fakeable category is needed in order to apply data-to-MC corrections for the probability with which jets passing the fakeable τ_h identification criteria pass the tight τ_h identification criteria, as detailed in [101].

4.2.5. Event-level variables

The $\not{E}_T$ (missing transverse energy or MET in short) and its related variable $\not{H}_T$ (also denoted by MHT) are computed from PF objects as described in appendix B.2. A linear discriminator, defined by

$$\not{E}_T^{\mathrm{LD}} = 0.00397 \cdot \not{E}_T + 0.00265 \cdot \not{H}_T\,, \tag{4.2}$$

has been developed by the $t\bar{t}h$ multilepton team, which helps to reject $Z+$jets events [102]. We use the recommended WP of $\not{E}_T^{\mathrm{LD}} > 0.2$ for the purpose of rejecting the $Z+$jets background.

4.3. Event selection

To reiterate, we want to probe the Higgs-to-top coupling by measuring the $t\bar{t}h$ production cross section in events in which the W boson from the top decays leptonically and the Higgs boson decays into a τ pair. By requiring one of the τ to decay leptonically and the other τ to decay hadronically, the final signature is expected to contain three leptons, one τ_h and two b-jets as shown in fig. 4.1. The neutrinos cannot be reconstructed as individual particles, but the vectorial sum of their transverse momenta can be inferred from the measured $\not{E}_T$. The cuts utilized in the event selection are dictated by the final signature of the signal $t\bar{t}h$ process, the geometry of the CMS detector, the quality of object reconstruction and identification, the kinematics of the dominant backgrounds contaminating the signal region, and the trigger paths used in the HLT.

First we perform a preselection meaning that we initially consider only those events that contain at least two jets (while ignoring their b-tagging CSV scores), three or more leptons passing loose preselection criteria and at least one τ_h passing the tight selection criteria listed in tables D.6, D.11, D.13 and D.14. This is followed by overlap removal: a set of objects are removed w.r.t other objects if they fall within predefined cone size around an object of higher priority. Muons are given the highest priority and therefore skip the cleaning step; electrons are cleaned w.r.t muons if they fall within a cone of size 0.3 around a muon; τ_h are cleaned w.r.t cleaned electrons and preselected muons using cone of size 0.3; finally, jets that fall within a cone of size 0.4 around muons, electrons or τ_h are excluded from further analysis. The overlap removal reassures that we do not

double-count the same particle as different types of objects. The removal of τ_h that overlaps with electrons or muons which pass the preselection criteria has the additional effect that it avoids the electrons or muons to be misidentified as τ_h if the electrons or muons fail the fakeable and tight selection criteria.

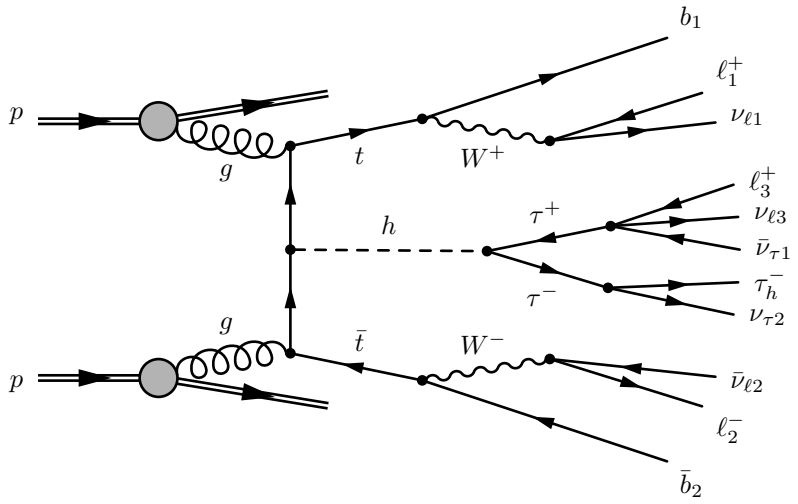


Figure 4.1: Feynman diagram for $t\bar{t}h$ process with $h \to \tau^+\tau^-$ decay mode corresponding to $3\ell + 1\tau_h$ channel.

The final event selection requires the electrons, muons and τ_h to pass tight cuts. Selected events are required to contain exactly three leptons passing tight selection criteria, at least one τ_h, and two or more jets. Either two of the jets need to pass loose WP of the CSV b-tagging algorithm or one jet passes the medium WP. The leading/subleading/third lepton must have p_T above[3] 25/10/10 GeV. In case the subleading lepton is an electron, it is required to satisfy $p_T > 15$ GeV. The charges of the three leptons (three leptons plus τ_h) must add up to ± 1 (zero). An event is vetoed if a pair of loose leptons have invariant mass lower[4] than 12 GeV, or if there exists a same-flavor opposite sign (SFOS) lepton pair passing the tight selection with invariant mass within 10 GeV to the Z boson mass. The requirement on $\not{E}_T^{\text{LD}}$ (eq. 4.2) is tightened to $\not{E}_T^{\text{LD}} > 0.3$ in case of an SFOS lepton pair because of higher possibility for Z boson contamination, or omitted altogether if the event has at least four jets since Z background is smaller at such high jet multiplicity.

The HLT paths used in this analysis are summarized in table D.15 which also lists event categories and primary datasets (PDs) where the trigger bits are applied. Since different triggers have recorded the same data (see table D.4), an elaborate logic has been implemented to avoid double-counting of the events in different PDs. For this reason we had to assign the triggers a priority so that lower priority trigger is ignored if the trigger of higher priority has already fired. The event selection is performed on both data and MC samples.

[3] The (sub)leading p_T cuts exceed nominal tight cuts because of the higher trigger thresholds used to collect the data.

[4] This phase space region is not well modeled by the MC simulations. The cut is applied to data as well because MC and data samples should be treated on equal grounds. This cut is also effective against virtual photon conversions.

4.3.1. Data-to-MC corrections

The simulation, reconstruction and identification have intrinsic efficiencies which need to be taken into account in order for the MC samples to match the data as closely as possible. Every MC event that passes the selection is given a total weight,

$$w_{\text{tot}} = w_{\text{lumi}} \times w_{\text{PU}} \times w_{\text{trig}} \times w_{\ell} \times w_{\tau h} \times w_{b\text{-tag}} \,, \tag{4.3}$$

where the weight w_{lumi} is used to normalize MC samples to integrated luminosity L, given its cross section σ and number of simulated events N_{event}:

$$w_{\text{lumi}} = \frac{\sigma \cdot L}{N_{\text{event}}} \,. \tag{4.4}$$

The PU weight w_{PU} accounts for differences in the number of PU interactions at given luminosity, which is also reflected in the distribution of PVs.

The trigger weights are parametrized as function of p_T of the two leading leptons passing the event selection:

$$w_{\text{trig}} = \frac{\varepsilon_{\text{data}}(p_T^{(1)}, p_T^{(2)})}{\varepsilon_{\text{MC}}(p_T^{(1)}, p_T^{(2)})} \,. \tag{4.5}$$

The trigger efficiencies in data ($\varepsilon_{\text{data}}$) and MC ($\varepsilon_{\text{MC}}$) are provided by the $t\bar{t}h$ multilepton analysis [59]. They are measured from events recorded with MET trigger because it is uncorrelated with the lepton triggers.

The lepton identification efficiency ε_{ℓ} is split into two parts for both data and MC,

$$\varepsilon_{\ell}^{X} = \varepsilon_{\ell X}^{\text{reco}\to\text{loose}} \times \varepsilon_{\ell X}^{\text{loose}\to\text{tight}} \,, \quad X \in \{\text{data, MC}\} \,, \tag{4.6}$$

where the first term represents the efficiency for reconstructed lepton to pass the loose selection, and the second term the efficiency for a loose lepton to pass the tight selection criteria. The efficiencies are different for electrons and muons, and are parametrized as function of p_T and η of the lepton. They have been measured by the $t\bar{t}h$ multilepton group in both data and MC with the tag-and-probe technique [103], using $Z \to \ell^+\ell^-$ events [59]. The event weight associated with the lepton identification is expressed by the ratio of the efficiency in data to the efficiency in MC:

$$w_{\ell} = \frac{\varepsilon_{\ell}^{\text{data}}}{\varepsilon_{\ell}^{\text{MC}}} \,. \tag{4.7}$$

The efficiencies of τ_h identification are provided by the Tau POG, who determined the exact figures via the tag-and-probe method using $Z/\gamma^* \to \tau^+\tau^-$ and $W \to \tau\nu_{\tau}$ events [99, 104]. Following their recommendation, we apply a weight of $w_{\tau h} = 0.99$ to events where the reconstructed τ_h matches a hadronic τ decay product on generator level, i.e. is a genuine τ_h and not a fake.

The b-tagging weights $w_{b\text{-tag}}$ are computed by the BTV POG [97]. These weights are determined by matching the distribution of the CSV b-tagging discriminator in MC to the distribution of the discriminator in data in bins of jet p_T and η for different types of jets [105]. The event weight $w_{b\text{-tag}}$ of the CSV b-tagging discriminator is obtained taking product of the individual b-tagging weights of each jet.

4.3.2. Background estimation

Even though the MC samples provide a good estimation for „irreducible" background sources such as $t\bar{t}W$, $t\bar{t}Z$ and WZ + jets in the signal region (SR), there still remains a possibility for jets and leptons in $t\bar{t}$ + jets events to take such kinematic configuration that the leptons and jets are reconstructed as prompt leptons or τ_h, thereby contaminating the SR. It is not feasible to model such backgrounds solely from MC simulations because this would require huge samples[5], as well as detailed understanding of detector and showering effects. In general, such „reducible" fake backgrounds where non-prompt leptons (i.e. hadrons or leptons arising from quarks or gluons) are misidentified as prompt leptons can be estimated from data by the means of data-driven fake factor (FF) method.

The FF technique relates events containing non-prompt leptons to events containing lepton candidates passing fakeable or tight cuts via fake factors $F_i = f_i/(1 - f_i)$. The fake rate (FR) f_i is computed for the i-th fakeable non-prompt lepton to pass the tight cuts from its p_T, $|\eta|$ and flavor. The FRs are measured from one control region (called „measurement region") but applied to another („application region") with the requirement that the non-prompt to prompt content in both regions are similar so that the method would not introduce any biases.

The fake background due to non-prompt leptons faking prompt ones is obtained from the application region, in which all selected leptons are required to pass the fakeable selection. To avoid overlap with the SR, the events are vetoed if all its selected fakeable leptons pass the tight selection. An event is weighed with FFs F_i, F_iF_j $(i \neq j)$ or $F_1F_2F_3$ if the i-th, i-th and j-th, or all fakeable leptons fail the tight cuts. The final event yield for the fake background is obtained by summing all events according to the prescription (see appendix C.1 for derivation)

$$N_{ppp}^{\text{bkg}} = \sum_{fpp} F_1 + \sum_{pfp} F_2 + \sum_{ppf} F_3 - \sum_{ffp} F_1F_2 - \sum_{fpf} F_1F_3 - \sum_{pff} F_2F_3 + \sum_{fff} F_1F_2F_3\,. \tag{4.8}$$

In this analysis we estimate the fake background due to non-prompt leptons only, meaning that τ_h selection is still kept tight. The FRs have already been measured by the $t\bar{t}h$ analysis team [60]. As for the application region, we use data events and weigh them according to (4.8). Because the data in the measurement region includes contributions from prompt leptons, we subtract MC events with prompt leptons that enter the application region, since the fake background needs to be estimated solely from non-prompt sources. Otherwise we would overestimate the fake background in the SR. The subtraction procedure is illustrated in fig. 4.2.

We do not include the τ_h in the fake background estimation procedure, in order not to lose sensitivity of the analysis. It turns out that in about 30% of $t\bar{t}h$ events selected in the SR the reconstructed τ_h is due to a misidentified quark or gluon jet. In case the τ_h were included in the fake background estimation procedure, about 30% of $t\bar{t}h$ signal events would be considered as „fake background" and not as „signal", thus significantly reducing the precision with which the $t\bar{t}h$ cross section can be measured. Instead of estimating the backgrounds with three prompt leptons and a fake τ_h from data, we apply data-to-MC

[5] Although the probability for multijet event entering the SR is low, its production cross section is very large. We cannot simply scale the event yields to nominal cross section because of the low statistics the MC samples currently provide.

correction to events that are selected in the SR and contain τ_h which do not match to a genuine τ_h, an electron or a muon on generator level. The data-to-MC corrections are determined from a $t\bar{t}$ + jets control region and are parametrized by the p_T and η of the τ_h candidate [101]. The corrections are fitted by a function linear in p_T. The uncertainty on the parameters of the linear function obtained from the fit are considered as systematic uncertainties.

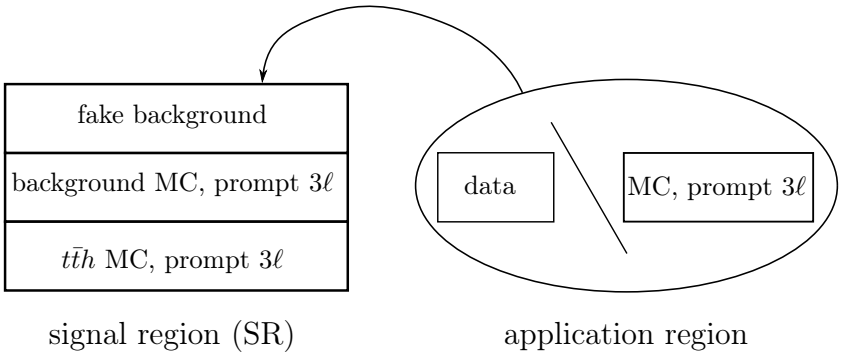


Figure 4.2: Composition of the SR (not to scale). The application region is defined by requiring all leptons to pass the fakeable selection but vetoing the events in which all leptons pass the tight selection criteria. The fake background estimate in the SR is computed by weighing all MC events with prompt leptons and the data with the FFs according to (4.8), followed by subtraction of the former from the latter.

4.4. Signal extraction

Despite the effort invested to optimize the event selection to select the $t\bar{t}h$ signal, but to remove backgrounds, the SR still contains a mixture of signal and background events. There is a tradeoff between tightening the event selection and maintaining the same statistical sensitivity of the analysis: tighter cuts result in lower event yields in both signal and background, as well as in data, which in turn increase the statistical uncertainties associated with the results. In other words, it would be harder to tell the difference between statistical fluctuations and actual discovery if the final event counts are too low.

However, we still have distributions of various kinematic variables at hand, which we can combine into a discriminant, that can be used to separate the $t\bar{t}h$ signal from backgrounds. The separation is not perfect, meaning that the two distributions (corresponding to signal and background) would still overlap. It is possible to mitigate the issue by means of a maximum likelihood fit (described in appendix C.2).

Additionally, there are different sources of systematic uncertainties arising from imperfect MC simulation, uncertainties of the data-to-MC corrections measured in data and on the data-driven fake background estimate. When we re-run the analysis with some variables shifted from their nominal values (e.g. applying larger JEC which corresponds to nominal correction plus one standard deviation times its uncertainty), it may change shape (if the correction under question depends on local kinematic properties) and/or normalization of the final discriminating variable. The maximum likelihood fitting model provides a handle to treat the aforementioned difficulties.

4.4.1. Systematic uncertainties

Systematic uncertainties never affect the data in the SR, only the signal and/or backgrounds that are compared to the data. There are two main sources of systematic uncertainties: experimental and theoretical, both of which are summarized in table D.16. The experimental sources are:

- 2.6% uncertainty on the measurement of the integrated luminosity [106], which affects only the normalization of MC samples but not fake background that is determined from the data;
- 3% uncertainty on the trigger efficiency [60];
- uncertainties on the lepton and τ_h identification efficiency: 5% for τ_h [104], and 2% (4%) uncertainty on the efficiency for leptons to pass loose (tight) selection criteria [60];
- uncertainties on the rate with which quark and gluon jets are misidentified as τ_h. The uncertainty is propagated to the SR by varying the parameters of the linear fit to the data-to-MC corrections, as described at the end of section 4.3.2 [5, 101];
- uncertainties on the FR measurements affect both the shape and the normalization of the fake background estimate, and are evaluated for electrons and muons separately depending on their p_T and η, as proposed by the $t\bar{t}h$ multilepton group [60];
- uncertainties on the JEC, which is applied by shifting the energy of the jets up or down before re-running the analysis (provided by the JetMET POG) [95, 96];
- uncertainties on the b-tagging efficiency provided by the BTV POG [97, 105];
- 3% uncertainty on the τ_h energy scale [104]. The systematics is similar to that of JEC: the energy of the τ_h is shifted by 3% up or down before re-evaluating the event selection.

Theoretical uncertainties on the inclusive cross section are obtained by varying the factorization and renormalization scales μ_F and μ_R, the PDFs and strong coupling constant α_s. These uncertainties (listed in table D.16) have been computed for $t\bar{t}h$, $t\bar{t}W$ and $t\bar{t}Z$ processes by the LHC Higgs Cross Section Working Group (LHCHXSWG) [32, table 231][107]. For the processes $t\bar{t}h$, $t\bar{t}W$ and $t\bar{t}Z$, we also consider the variation of the renormalization and factorization scales[6] between $Q/2$ and $2Q$ as separate shape systematics. Here Q refers to the nominal energy scale, which equals half the sum of masses corresponding to the event (for instance, the nominal energy scale for $t\bar{t}h$ is $Q = m_t + m_h/2$ [108]). An uncertainty of 50% is assigned to the (minor) backgrounds arising from diboson, Z + jets and W + jets production.

[6] The renormalization and factorization scales refer to energy values used in the regularization to absorb ultraviolet and infrared divergences, respectively, in the perturbative expressions used to compute the cross section. Because these energy scales are artifacts of the computation and the physics should be independent of these scales, the evaluation of cross sections and differential shapes thereof at different scales yields an estimate on uncertainties for missing higher order contributions in the perturbative expansion.

4.4.2. Final discriminant

The $t\bar{t}h$ multilepton analysis group has developed BDT-based algorithms for separating the $t\bar{t}h$ signal from the irreducible $t\bar{t}W$ and $t\bar{t}Z$, as well as from the reducible $t\bar{t}$ + jets backgrounds [60]. The training of the BDT separating the signal from the irreducible ($t\bar{t}V$) and from the reducible ($t\bar{t}$) backgrounds uses the following observables as inputs:

- maximum $|\eta|$ of the two leading leptons (both BDTs);
- number of preselected jets in the event (both BDTs);
- angular separation (given by B.13) between the selected leptons and jets closest to them (both BDTs);
- transverse mass of the leading lepton (given by B.9; both BDTs);
- missing transverse energy (given by B.11; both BDTs);
- MHT (given by B.12; $t\bar{t}$ BDT) or p_T of the leading lepton ($t\bar{t}V$ BDT);
- arithmetic average of angular separations between jet pairs ($t\bar{t}$ BDT) or p_T of the third lepton ($t\bar{t}V$ BDT).

The output of the two BDTs ranges from -1 to $+1$ and are used to construct two-dimensional plane ($[-1,+1]\times[-1,+1]$), which is divided into five regions (as shown in table 4.1). Each region is assigned a different value of the final discriminant D_{MVA} that is used for the signal extraction. The distribution of both BDTs and the final discriminant are shown in figures 4.3, 4.4 and 4.5 for data, signal and background. The final event yields in the SR are presented in table 4.2.

Final discriminant D_{MVA}	$t\bar{t}$ BDT	$t\bar{t}V$ BDT
1	$(-1,-0.3]$	$(-1,1]$
2	$(-0.3,0.3]$	$(-1,0.25]$
3	$(-0.3,0.3]$	$(0.25,1]$
4	$(0.3,1]$	$(-1,0.25]$
5	$(0.3,1]$	$(0.25,1]$

Table 4.1: Binning of $t\bar{t}h$ vs $t\bar{t}$ and $t\bar{t}h$ vs $t\bar{t}V$ discriminants.

Process	Event counts in the SR
$t\bar{t}h$, $h \to \tau^+\tau^-$	1.01 ± 0.65
$t\bar{t}h$, $h \to W^+W^-$	0.63 ± 0.29
$t\bar{t}h$, $h \to ZZ$	0.09 ± 0.04
$t\bar{t}Z$	3.78 ± 0.62
$t\bar{t}W$	0.24 ± 0.05
electroweak	0.32 ± 0.05
fake background	1.07 ± 0.34
other background	0.24 ± 0.08
total expected background	5.65 ± 0.85
total SM expectation	7.38 ± 1.10
observed data	7

Table 4.2: Event yields in the SR. Quoted uncertainties include both statistical and systematic sources.

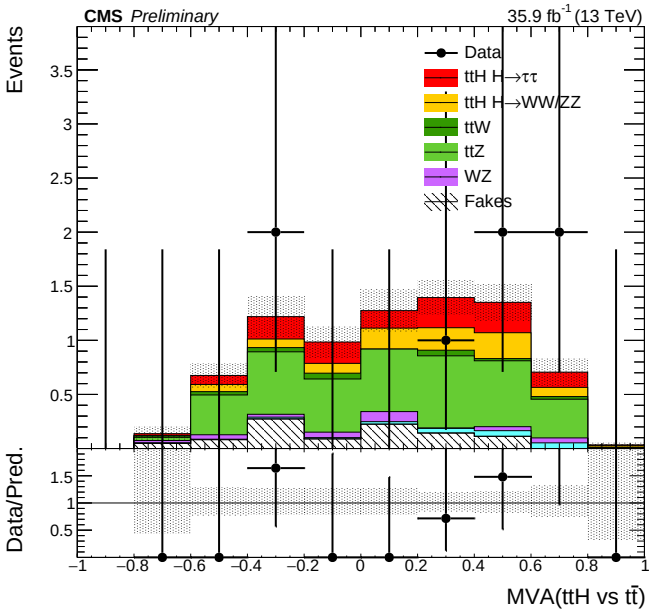


Figure 4.3: Distribution in the output of the $t\bar{t}$ BDT that is used for the signal extraction of the events selected in the SR. The area shaded in gray indicates the sum of statistical and systematic uncertainties on the sum of $t\bar{t}h$ signal plus backgrounds. Statistical uncertainties on data are shown with vertical lines.

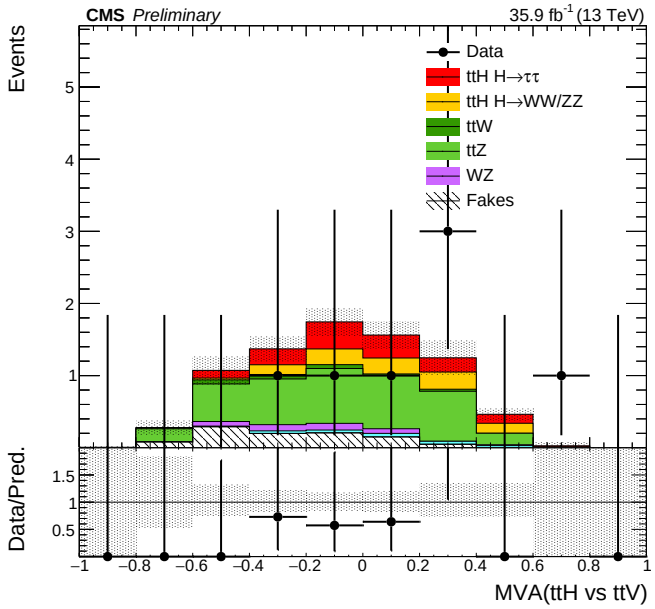


Figure 4.4: Distribution in the output of the $t\bar{t}V$ BDT that is used for the signal extraction of the events selected in the SR. The area shaded in gray indicates the sum of statistical and systematic uncertainties on the sum of $t\bar{t}h$ signal plus backgrounds. Statistical uncertainties on data are shown with vertical lines.

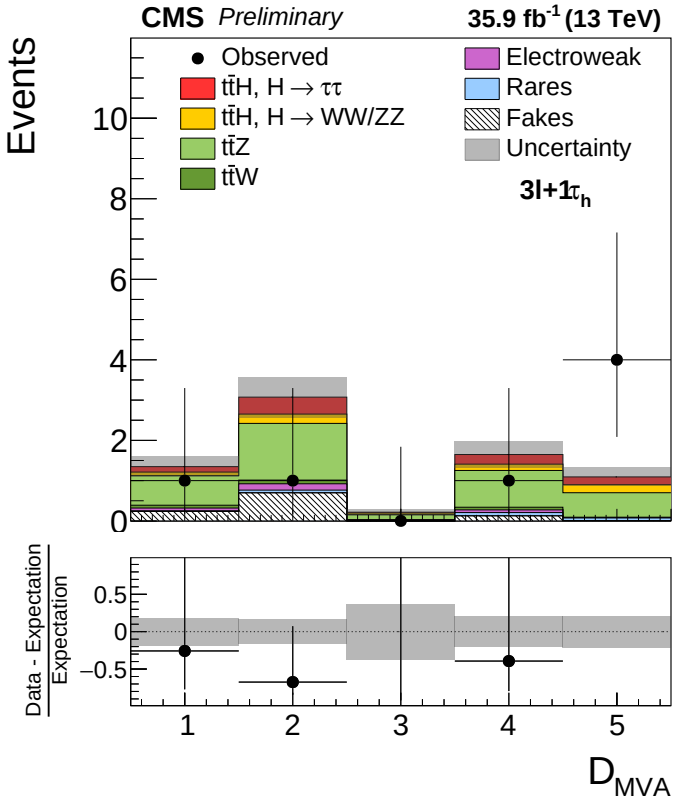


Figure 4.5: Distribution in the final discriminant that is used for the signal extraction of the events selected in the SR. The area shaded in gray indicates the sum of statistical and systematic uncertainties on the sum of $t\bar{t}h$ signal plus backgrounds. Statistical uncertainties on data are shown with vertical lines.

4.4.3. Matrix element method

The matrix element method (MEM) is based on computing the probability $\mathcal{P}_i(\hat{\mathbf{p}})$ for a process i to occur inside the detector, given the experimental evidence $\hat{\mathbf{p}}$ [109, 110]. The probability is evaluated by computing the integral

$$\mathcal{P}_i(\hat{\mathbf{p}}) = \frac{1}{\sigma_i} \sum_{\text{perm}} \frac{1}{N} \int \mathrm{d}x_a \, \mathrm{d}x_b \, \mathrm{d}\Phi_n \, \frac{f(x_a;\, Q_i) f(x_b;\, Q_i)}{2 x_a x_b s} \cdot (2\pi)^4 \cdot |\mathcal{M}_i(\mathbf{p})|^2 \, W(\hat{\mathbf{p}}|\mathbf{p}) \,, \quad (4.9)$$

where the sum runs over different permutations of associating parton-level objects with the objects reconstructed in the detector, explained in more detail below. The constant $\frac{1}{N}$ is a normalization factor associated with these permutations. The integration is performed over the Bjorken variables x_a and x_b, which correspond to the fraction of proton momentum that is carried by the two partons initiating the process i, and over the Lorentz-invariant phase space element (LIPS)

$$\mathrm{d}\Phi_n = \delta^{(4)} \left(x_a P_a + x_b P_b - \sum_{k=1}^{n} p_k \right) \prod_{k=1}^{n} \frac{\mathrm{d}^3 \mathbf{p}_k}{(2\pi)^3 2 E_k} \,, \quad (4.10)$$

which corresponds to the phase space of the n final-state particles of 4-momentum $p_k = (E_k,\, \mathbf{p}_k)$ that are produced by the process i. The squared matrix element of the process is given by $|\mathcal{M}_i(\mathbf{p})|^2$. The PDFs f are evaluated at the scale Q_i. The transfer functions $W(\hat{\mathbf{p}}|\mathbf{p})$ represent the probability densities for observing measured momenta $\hat{\mathbf{p}}$, given a phase space point $\mathbf{p}$. The probability (4.9) is normalized to the total cross section σ_i of the process i, and is computed at the center-of-mass energy $\sqrt{s}$. The measured momenta $\hat{\mathbf{p}}$ and the true momenta $\mathbf{p}$ are defined in the laboratory frame.

The transfer functions (TFs) map the true (parton-level) momenta $\mathbf{p}$ to the momenta $\hat{\mathbf{p}}$ measured by the detector. By requiring the TFs to be normalized to unity,

$$\int \mathrm{d}\hat{\mathbf{p}} \, W(\hat{\mathbf{p}}|\mathbf{p}) = 1 \quad \forall \, \mathbf{p} \,, \quad (4.11)$$

the overall probability (4.9) is then normalized to unity as well:

$$\int \mathrm{d}\hat{\mathbf{p}} \, \mathcal{P}_i(\hat{\mathbf{p}}) = 1 \,, \quad (4.12)$$

since the total cross section of the process i is given by

$$\sigma_i = \int \mathrm{d}x_a \, \mathrm{d}x_b \, \mathrm{d}\Phi_n \, \frac{f(x_a;\, Q_i) f(x_b;\, Q_i)}{2 x_a x_b s} \cdot (2\pi)^4 \cdot |\mathcal{M}_i(\mathbf{p})|^2 \, W(\hat{\mathbf{p}}|\mathbf{p}) \,. \quad (4.13)$$

The normalization factor $\frac{1}{N}$ is chosen such that $\sum_{\text{perm}} \frac{1}{N} = 1$. As explained in section 2.3, the partons initiating the hard process are typically gluons, which carry the fraction x_a and x_b of the colliding protons which move with the following 4-momenta inside the detector:

$$P_{a,b} = \left(\frac{\sqrt{s}}{2},\, 0,\, 0,\, \pm \frac{\sqrt{s}}{2} \right) \,. \quad (4.14)$$

For this reason we only consider the gluon PDF and not the quark PDFs in (4.9).

The Neyman-Pearson lemma states that the likelihood ratio

$$\Lambda(\hat{\mathbf{p}}) = \frac{\mathcal{P}_0(\hat{\mathbf{p}})}{\mathcal{P}_0(\hat{\mathbf{p}}) + \mathcal{P}_1(\hat{\mathbf{p}})} \in [0, 1] \tag{4.15}$$

provides the best separation between the signal hypothesis ($i = 0$) and the background hypothesis ($i = 1$). The ratio given by (4.15) is computed for all events: simulated signal and background events, as well as for data. Provided that the MEM is implemented properly, we expect to obtain higher values of $\Lambda(\hat{\mathbf{p}})$ for signal events compared to background events.

Motivated by the event yields shown in table 4.2, I implemented the MEM for two hypotheses: for the $t\bar{t}h$ signal process and for the $t\bar{t}Z$ background process. Generator-level matching of the selected events in the SR indicates that majority of these events correspond to the process in which the Higgs (or Z) boson decays into a pair of τ leptons, from which one decays leptonically and the other hadronically; the W bosons from the pair of top quarks decay leptonically, giving rise to three leptons and one τ_h in the final state. The Feynman diagram for the signal ($t\bar{t}h$) hypothesis is shown in fig. 4.1; the diagram corresponding to the background ($t\bar{t}Z$) hypothesis is obtained by replacing h with Z in this diagram.

The final state contains a total of $n = 11$ particles: 3 leptons, a τ_h, 2 b-jets and 5 neutrinos. Therefore, the LIPS $\mathrm{d}\Phi_n$ corresponds to 44 integration variables. We can reduce the number of integration variables by exploiting the recursive structure of the LIPS, and by requiring the intermediate particles (top quarks, W bosons, the Higgs boson or, alternatively, the Z boson) to be on their mass shells. This is achieved by approximating the propagators inside the matrix element by δ-functions, given by eq. (A.31). With these assumptions, the total number of integration variables is reduced to 8. These d.o.f correspond to the unmeasured momenta of the neutrinos. The integration variables are chosen such that the correlation between any two integration variables is very small. The reduction of the phase space integral to 8 dimensions relies on the results of [111, 112]. The integration variables are: the azimuthal angles and the cosines of the polar angles of the two neutrinos originating from the decays of the top quarks ($\phi_{1,2}$ and $\cos\theta_{1,2}$); the energy fraction carried by the τ_h (z_1); the angles $\phi^{\mathrm{inv}}(\tau_h, \nu)$ and $\phi^{\mathrm{inv}}(\ell, 2\nu)$ in the τ decay system, which are visualized in fig. 4.6; the squared mass of the neutrino pair produced in the leptonic τ decay (m^2_{inv}). The integration limits and the correlation matrix of these integration variables are given in tables D.17 and D.18, respectively. The integration variables are defined in the laboratory frame.

The integral given by (4.9) is evaluated numerically, using a custom implementation of the Markov Chain integration method [113, 114]. The integration variables and integration limits are given in table D.17. We have compared the performance of the Markov Chain integrator for the alternative algorithms VEGAS [115] and VAMP [116], which are commonly used in high energy physics applications, and confirmed that our custom implementation of the Markov Chain integrator yields similar numerical values, while having the advantage of reducing the computing time. The implementation of our MEM code is made available at [4].

At every integration step we reconstruct the 4-momenta of all particles in the final state (except for the two neutrinos originating from the leptonic τ decay – their 4-momenta is summed by construction), and use these momenta to evaluate the matrix element $|\mathcal{M}|^2$

in the frame of reference in which the total transverse momentum of the reconstructed particles is zero. This choice of reference frame is necessary since we use a LO matrix element generated by MadGraph5_aMC@NLO. A complication arises from the fact that this program does not model the decays of τ leptons. In order to evaluate the matrix element, we need to reconstruct the 4-momenta of the two τ leptons and use them as an input when evaluating $|\mathcal{M}|^2$.

$$
\begin{aligned}
e_z &= \frac{\mathbf{p}^{\text{vis}}}{|\mathbf{p}^{\text{vis}}|} \\
e_y &= \frac{e_z \times (0,0,1)}{|e_z \times (0,0,1)|} \\
e_x &= e_y \times e_z
\end{aligned}
$$

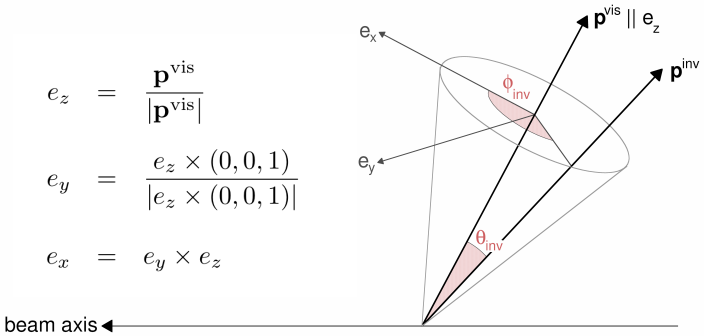


Figure 4.6: Illustration of the angles $\phi^{\text{inv}}(\tau_h, \nu)$ and $\phi^{\text{inv}}(\ell, 2\nu)$ that we use to model the hadronic, as well as the leptonic τ decays [112, Fig. 4]. The 3-momentum $\mathbf{p}^{\text{vis}}$ corresponds to the „visible" τ decay product, which is either a lepton or a τ_h; the 3-momentum $\mathbf{p}^{\text{inv}}$ corresponds to the „invisible" τ decay product, which is either a single neutrino in case of a hadronic τ decay, or the vectorial sum of the two neutrinos produced in case of a leptonic τ decay. The local coordinate system is built around the visible momentum vector such that the z-axis coincides with the direction of the visible momentum; the y-axis is perpendicular to the beam axis and to the z-axis; the x-axis is perpendicular to the local y-z plane. The symbol θ_{inv} denotes the opening angle between the visible and the invisible τ decay products. The rotation angle ϕ^{inv} is measured on the y-x plane and is chosen such that the case $\phi_{\text{inv}} = 0$ corresponds to the configuration in which $\mathbf{p}^{\text{inv}}$ lies on the plane spanned by the beam axis and $\mathbf{p}^{\text{vis}}$.

Concerning the permutations: the final state always contains two leptons of the same charge and we do not know as which of these leptons originates from a W decay and which originates from a τ decay, so we need to evaluate the matrix element $|\mathcal{M}|^2$ for the two permutations of leptons that have the same charge. A similar association problem occurs for the b-jets as well – the charge of the b-jets is not measured reliably by the detector and we therefore do not know which b-jet is associated with which top quark. So, we have to evaluate the matrix element also for the two permutations of b-jets. The event selection may even yield more than two reconstructed b-jets. In order to maximize the probability that the correct assignment of reconstructed b-jets to the b quarks produced in the top decays is included in the sum over the permutations, we consider permutations of the b quarks to up to three reconstructed b-jets, sorted by their CSV b-tagging score. Thus, we perform up to 12 evaluations of the matrix element by considering all possible associations between reconstructed and parton-level objects, in order to obtain the final probability of an event to match a given hypothesis.

The reconstructed 4-momenta of the electrons and the muons, as well as the 4-momenta of the hadronic τ decay products are used as an input to the MEM. In other words, we assume that the 4-momenta of electrons, muons and of τ_h are measured perfectly by the detector. The reconstructed 4-momenta of the b-jets, however, are recalculated from the requirement that the mass of a b-jet is equal to the mass of a b-quark ($m_b = 4.7\,\text{GeV}$), while assuming the conservation of momentum in the decay $t \to Wb$. The direction of b-jet,

specified by the angles $\hat{\eta}_b$ and $\hat{\phi}_b$, is kept constant, reflecting the fact that the detector has good angular resolution, but poor energy resolution for the jets. The recalculated jet energy E_b (i.e. the „true" parton-level energy) and the reconstructed energy $\hat{E}_b$ (measured in the detector) of a b-jet enter to the TF:

$$W_{Eb}(\hat{E}_b|E_b) = \frac{f_b(E_b,\,\hat{\eta}_b)}{\sqrt{2\pi}\sigma_b(E_b,\,\hat{\eta}_b)} \exp\left[-\frac{(\hat{E}_b - \mu_b(E_b,\,\hat{\eta}_b))}{2\sigma_b^2(E_b,\,\hat{\eta}_u)}\right] + \tag{4.16}$$

$$+ \frac{(1 - f_b(E_b,\,\hat{\eta}_b))}{\sqrt{2\pi}\sigma'_b(E_b,\,\hat{\eta}_b)} \exp\left[-\frac{(\hat{E}_b - \mu'_b(E_b,\,\hat{\eta}_b))}{2\sigma'^2_b(E_b,\,\hat{\eta}_u)}\right] . \tag{4.17}$$

The parameters f_b, σ_b, μ_b, σ'_b and μ'_b of the two superimposed Gaussian distributions are given in [111]. The function W_{Eb} determines the probability for a b quark of energy E_b to be reconstructed as a b-jet of energy $\hat{E}_b$ in the detector. The TF of b quarks is shown in fig. 4.7. We also use another TF to quantify the resolution on $\not{E}_T$. More specifically, we compare the vectorial sum of the true transverse momenta of the neutrinos ($\not{\mathbf{E}}_T$) to the measured value of the missing transverse energy ($\hat{\not{\mathbf{E}}}_T$):

$$W(\not{\mathbf{E}}_T|\hat{\not{\mathbf{E}}}_T) = \frac{1}{2\pi\sqrt{|V_{xy}|}} \exp\left[-\frac{1}{2}\begin{pmatrix}\Delta\not{E}_x \\ \Delta\not{E}_y\end{pmatrix}^T V_{xy}^{-1} \begin{pmatrix}\Delta\not{E}_x \\ \Delta\not{E}_y\end{pmatrix}\right] , \tag{4.18}$$

where the differences are given by $\Delta\not{\mathbf{E}}_T \equiv (\Delta\not{E}_x,\, \Delta\not{E}_y) = \hat{\not{\mathbf{E}}}_T - \not{\mathbf{E}}_T$, and the symbol V_{xy} refers to the covariance matrix of $\hat{\not{\mathbf{E}}}_T$, estimated by the algorithm described in [117]. The Bjorken variables x_a and x_b are determined by the relation

$$x_{a,b} = \frac{1}{\sqrt{s}}(E^{\text{tot}} \pm p_z^{\text{tot}}) . \tag{4.19}$$

where E^{tot} and p_z^{tot} refer to the total energy and longitudinal momentum of all particles in the final state:

$$E^{\text{tot}} = \sum_{k=1}^{n} E_k \qquad \text{and} \qquad p_z^{\text{tot}} = \sum_{k=1}^{n} p_z^k . \tag{4.20}$$

The level of separation between the $t\bar{t}h$ signal and the $t\bar{t}Z$ background that is achieved by the current implementation of the MEM is visualized in fig. 4.8. For the purpose of this plot, the reconstructed objects in the events are required to be matched to their corresponding generator-level objects. We used five different settings to evaluate the performance of the MEM (the corresponding line colors in fig. 4.8 are written in parentheses):

- use true 4-momenta for all objects, evaluate all integration variables from the true MC information (i.e. no integration is performed; purple);
- use true 4-momenta for all objects, integrate over hadronic energy fraction z_1, while all other integration variables are computed from the true MC information (i.e. integrand is 1-dimensional; blue);

- use true 4-momenta for all objects, integrate over all integration variables, except for the hadronic energy fraction z_1 (i.e. integrand is 7-dimensional; black);
- use true 4-momenta for all objects, integrate over all integration variables (i.e. integrand is 8-dimensional; green);
- use reconstructed 4-momenta for all objects, integrate over all integration variables (red).

The plot indicates that the performance of separating the $t\bar{t}h$ signal from the $t\bar{t}Z$ background is limited by the fact that the values of the integration variables are not optimal, even for a „perfect" detector (purple vs green curve). The performance is further reduced because of the experimental resolution of the detector (green vs red curve). These results are expected. The fact that the red curve still performs superior for random guessing (dashed line), it demonstrates that the MEM is viable.

Unfortunately, the current implementation of the MEM is not good enough to use it as a signal extraction variable, or as an input to a BDT. If we remove the requirement of the reconstructed objects to be matched to their generator-level objects, the performance of the MEM degrades so much that it does not perform much better than random guessing anymore. The method can be improved by including additional signal and background hypotheses that cover the case in which one or more reconstructed objects is a fake. For instance, there is a sizable contribution to the SR from events in which a Higgs boson decays into a pair of W bosons, which further decay leptonically, and one of the W bosons from the top quarks decays leptonically, while the other W boson decays into a τ lepton, which subsequently decays into hadrons (symbolically: $H \to WW \to \ell\ell$ + neutrinos, $t\bar{t} \to \ell\tau \to \ell\tau_h$ + neutrinos). These improvements need to be implemented in the MEM before the likelihood ratio computed by the MEM can be used as signal extraction variable in the analysis.

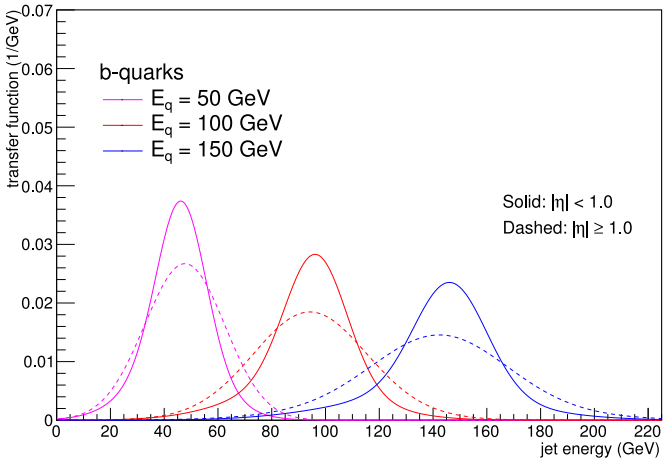


Figure 4.7: The b-jet energy TF for three values of true b quark energies. The plot is reproduced from [111, Figure 2].

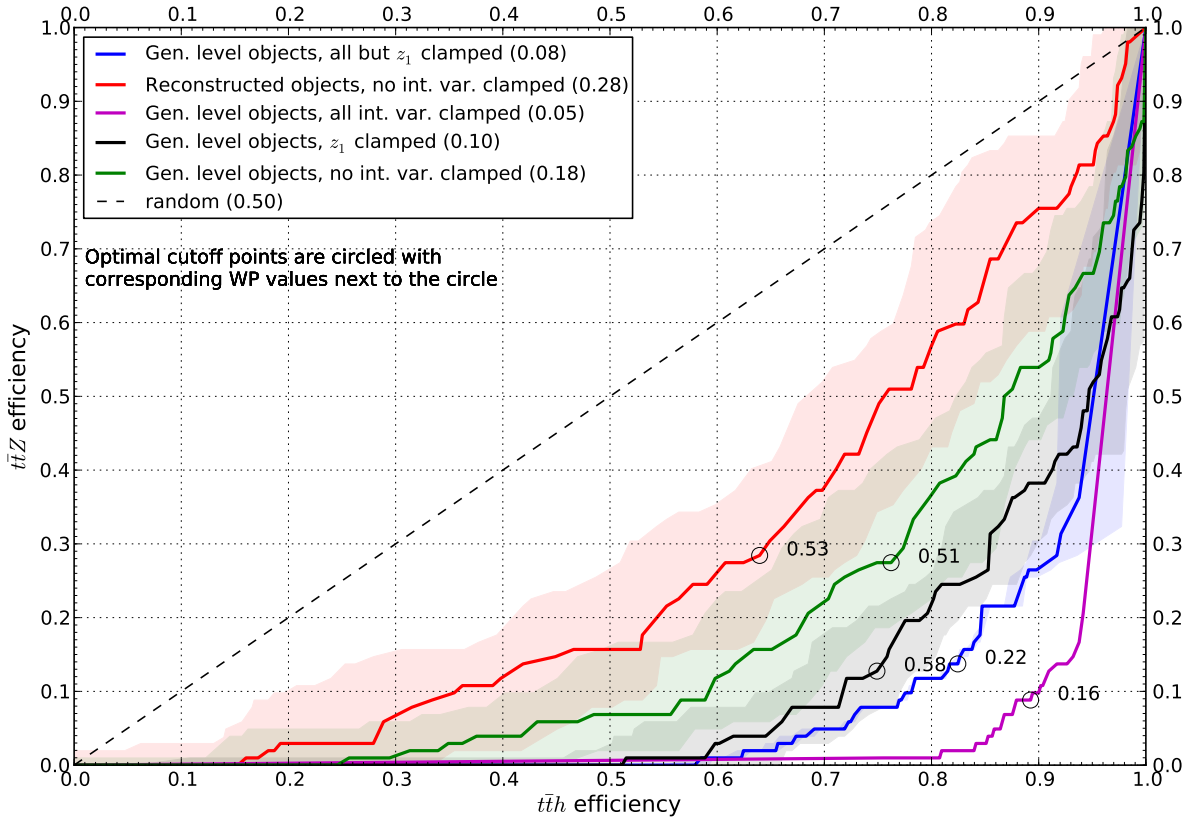


Figure 4.8: Receiver operator characteristic (ROC) curve of the MEM likelihood ratio $\Lambda(\hat{\mathbf{p}})$, given by (4.15), for the $t\bar{t}h$ signal hypothesis versus the $t\bar{t}Z$ background hypothesis. The MEM likelihood ratio is evaluated for $t\bar{t}h$ and $t\bar{t}Z$ events that are selected in the SR and in which all reconstructed objects are matched to their generator-level objects. A single point in the ROC curve is obtained by determining the fraction of simulated $t\bar{t}h$ signal and $t\bar{t}Z$ background events for which the likelihood ratio $\Lambda(\hat{\mathbf{p}})$ is greater than or equal to a given cut threshold. The full curve is drawn by varying the threshold between 0 and 1. The optimal cut on $\Lambda(\hat{\mathbf{p}})$ corresponds to the point on the ROC curve that is closest to the bottom right corner of the plot. The numbers quoted within parentheses in the legend quantify the area under the ROC curves (smaller values indicate better performance). The colored bands indicate errors in numerical integration.

4.4.4. Results

Given the shapes and event yields expected for the $t\bar{t}h$ signal and for different background contributions in the discriminant D_{MVA} shown in fig. 4.5, and the set of systematic uncertainties described in section 4.4.1, we can build a maximum likelihood (ML) model. The signal strength, given by (2.38), is taken as parameter of interest (POI). It is left unconstrained in the fit and is expected to take value of 0 for the background-only hypothesis and 1 for the signal-plus-background hypothesis in the SM. Systematic uncertainties are represented by nuisance parameters in the fit. The details of the ML fit are presented in appendix C.2. We just quote the final results here: the observed signal rate amounts to $\mu = 1.22^{+1.34}_{-1.00}$, while the expected value amounts to $\mu = 1.00^{+1.41}_{-1.06}$. The result corresponds to a measured $t\bar{t}h$ signal cross section of $\sigma(t\bar{t}h) = 618.7^{+679.5}_{-507.1}\,\text{fb}$, while the SM predicts $\sigma_{\text{SM}} = 507.1\,\text{fb}$ for a Higgs boson mass of $m_h = 125\,\text{GeV}$.

At the time of writing, there are ongoing efforts to combine the results of the $t\bar{t}h$ multilepton analysis with the analysis presented in this thesis. The multilepton analysis considered three final states without τ_h and measured a signal rate of $\mu = 1.78^{+0.60}_{-0.54}$ (for an expected value of $\mu = 1.00^{+0.46}_{-0.42}$). The results are interpreted as yielding a significance of 3.4σ (2.4σ to observe evidence for a $t\bar{t}h$ signal in the CMS data [60]. The $t\bar{t}h$, $h \to \tau^+\tau^-$ analysis presented in [5] includes two more final states with τ_h in addition to the $3\ell + 1\tau_h$ final state presented in this thesis, and measures a signal rate of $\mu = 0.72^{+0.62}_{-0.53}$ (for an expected value of $\mu = 1.00^{+0.67}_{-0.57}$) which is interpreted as observing evidence for $t\bar{t}h$ production with a significance of 1.4σ (for an expected significance of 1.8σ).

In conclusion, with the given amount of data, the results are in agreement with the SM expectation, but there is certainly room for improvement because the background-only hypothesis has not been completely ruled out, yet.

5. Summary

In this work, we probed the Higgs-to-top Yukawa coupling by measuring the $t\bar{t}h$ production cross section in events containing three electrons or muons and one hadronic τ decay product in final states (denoted by $3\ell + 1\tau_h$). The analysis has been performed on data recorded by the CMS detector in 2016 and it amounts to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. The data corresponds to proton-proton collisions at center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ with a bunch spacing of 25 ns.

The analysis presented in this thesis relied on the work done by other dedicated groups of experts in CMS. In particular, the Physics Object Groups (POGs) have developed and validated algorithms for reconstructing and identifying various physics objects that we used in this analysis: electrons, muons, jets and hadronic τ decay products. We followed their guidelines on how to use these objects.

The signal $t\bar{t}h$ process and irreducible background interactions are estimated from Monte Carlo (MC) simulations. The fake background due to misidentification of non-prompt leptons (i.e. leptons originating from quark or gluon decays) as prompt leptons (i.e. leptons originating from the decays of W boson, Z boson or τ lepton) in the signal region is extrapolated from the data via fake factor method.

The signal region is defined by cut-based event selection targeting the signal $t\bar{t}h$ events that correspond to $3\ell + 1\tau_h$ final state. However, the signal region still includes substantial event yields from the background processes, predominantly from the $t\bar{t}Z$ process. Differences between the MC simulations and the data have been corrected by appropriately weighing the selected MC events. The data-to-MC corrections as well as systematic uncertainties are provided by the POGs, the $t\bar{t}h$ multilepton analysis group and other CMS groups.

The experimental and theoretical sources of shape and yield systematics, as well as their uncertainties have been estimated. The final discriminant is constructed from two BDT discriminants optimized to separate the signal $t\bar{t}h$ events from the $t\bar{t}$ + jets, $t\bar{t}W$ and $t\bar{t}Z$ events. As per CMS standard, the signal extraction is performed by means of a maximum likelihood fit to the data in which the systematic uncertainties are treated as nuisance parameters.

The measured signal strength amounts to $\mu = 1.22^{+1.34}_{-1.00}$ times the $t\bar{t}h$ production cross section predicted by the Standard Model (SM). We conclude that our results are compatible with the SM predictions, although there is still much room for improvements on uncertainties of the measurement. The work continues with further development of the MEM.

Bibliography

[1] S. Chatrchyan et al., Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. **B716**, 30–61 (2012).

[2] G. Aad et al., Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. **B716**, 1–29 (2012).

[3] NICPB Tallinn group, Code and Python config files for $t\bar{t}h$, $h \to \tau\tau$ analysis with matrix element techniques, https://github.com/HEP-KBFI/tth-htt/, [Online, retrieved 27-April-2017], 2017.

[4] NICPB Tallinn group, Matrix element method for $t\bar{t}H$, $H \to \tau\tau$ channel, https://github.com/HEP-KBFI/tth-analyze-mem/, [Online, retrieved 24-May-2017], 2017.

[5] Search for the associated production of a Higgs boson with a top quark pair in final states with a τ lepton at $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-HIG-17-003 (CERN, Geneva, 2017).

[6] S. L. Glashow, Partial-symmetries of weak interactions, Nuclear Physics **22**, 579–588 (1961).

[7] S. Weinberg, A Model of Leptons, Physical Review Letters **19**, 1264–1266 (1967).

[8] A. Salam, Weak and Electromagnetic Interactions, in Elementary particle theory, Relativistic groups and analyticity, Proceedings of the Eighth Nobel Symposium, edited by N. Svartholm (1968), pp. 367–377.

[9] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Advantages of the Color Octet Gluon Picture, Phys. Lett. **B47**, 365–368 (1973).

[10] M. Gell-Mann, A Schematic Model of Baryons and Mesons, Phys. Lett. **8**, 214–215 (1964).

[11] D. J. Gross and F. Wilczek, Asymptotically Free Gauge Theories. I, Physical Review D **8**, 3633–3652 (1973).

[12] G. Zweig, An SU_3 model for strong interaction symmetry and its breaking; Version 2, 80 p (1964).

[13] F. Englert and R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, Physical Review Letters **13**, 321–323 (1964).

[14] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Global Conservation Laws and Massless Particles, Physical Review Letters **13**, 585–587 (1964).

[15] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Physical Review Letters **13**, 508–509 (1964).

[16] P. Higgs, Broken symmetries, massless particles and gauge fields, Physics Letters **12**, 132–133 (1964).

[17] P. W. Higgs, Spontaneous Symmetry Breakdown without Massless Bosons, Physical Review **145**, 1156–1163 (1966).

[18] T. W. B. Kibble, Symmetry Breaking in Non-Abelian Gauge Theories, Physical Review **155**, 1554–1561 (1967).

[19] G. Aad, B. Abbott, J. Abdallah, et al., Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ TeV and 8 TeV with the ATLAS and CMS Experiments, Physical Review Letters **114** (2015) 10.1103/physrevlett.114.191803.

[20] L. Alvarez-Gaumé and M. Vázquez-Mozo, *An Invitation to Quantum Field Theory*, Lecture Notes in Physics (Springer Berlin Heidelberg, 2011), ISBN: 9783642237270.

[21]H. E. Logan, TASI 2013 lectures on Higgs physics within and beyond the Standard Model, (2014).

[22]M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Advanced book classics (Avalon Publishing, 1995), ISBN: 9780201503975.

[23]M. Robinson, *Symmetry and the Standard Model: Mathematics and Particle Physics*, SpringerLink: Bücher (Springer New York, 2011), ISBN: 9781441982674.

[24]K. Olive, Review of Particle Physics, Chinese Physics C **40**, 100001 (2016).

[25]M. Froissart, Asymptotic Behavior and Subtractions in the Mandelstam Representation, Physical Review **123**, 1053–1057 (1961).

[26]M. Szleper, The Higgs boson and the physics of WW scattering before and after Higgs discovery, (2014).

[27]B. Mellado Garcia, P. Musella, M. Grazzini, and R. Harlander, CERN Report 4: Part I Standard Model Predictions, (2016).

[28]Off-shell Higgs signal strength measurement using high-mass $H \to ZZ \to 4\ell$ events at High Luminosity LHC, tech. rep. ATL-PHYS-PUB-2015-024 (CERN, Geneva, 2015).

[29]S. Heinemeyer, C. Mariotti, G. Passarino, R. Tanaka, et al., Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group, tech. rep. CERN-2013-004. (Geneva, 2013).

[30]P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn, Scalar Correlator at $\mathcal{O}(\alpha^4)$, Higgs Boson Decay into Bottom Quarks, and Bounds on the Light-Quark Masses, Physical Review Letters **96** (2006) 10.1103/physrevlett.96.012003.

[31]A. Dabelstein and W. Hollik, Electroweak corrections to the fermionic decay width of the standard Higgs boson, Zeitschrift für Physik C Particles and Fields **53**, 507–515 (1992).

[32]D. de Florian et al., Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, (2016).

[33]J. Campbell, Cross sections at 14, 33, and 100 TeV pp colliders, Snowmass: Seattle Energy Frontier Workshop, 2013.

[34]P. A. Baikov and K. G. Chetyrkin, Top Quark Mediated Higgs Boson Decay into Hadrons to Order α^5, Physical Review Letters **97** (2006) 10.1103/physrevlett.97.061803.

[35]U. Aglietti et al., Tevatron for LHC report: Higgs, in (2006).

[36]M. Inoue, R. Najima, T. Oka, and J. Saito, QCD corrections to two photon decay of the Higgs boson and its reverse process, Modern Physics Letters A **09**, 1189–1194 (1994).

[37]M. Spira, A. Djouadi, and P. Zerwas, QCD corrections to the $HZ\gamma$ coupling, Physics Letters B **276**, 350–353 (1992).

[38]G. Degrassi and F. Maltoni, Two-loop electroweak corrections to Higgs production at hadron colliders, Physics Letters B **600**, 255 –260, ISSN: 0370-2693 (2004).

[39]P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, Vector boson fusion at next-to-next-to-leading order in QCD: Standard model Higgs boson and beyond, Physical Review D **85** (2012) 10.1103/physrevd.85.035002.

[40]T. Figy, S. Palmer, and G. Weiglein, Higgs production via weak boson fusion in the standard model and the MSSM, Journal of High Energy Physics **2012**, 105, ISSN: 1029-8479 (2012).

[41]L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke, Gluon-induced Higgs-strahlung at next-to-leading order QCD, Journal of High Energy Physics **2013**, 78, ISSN: 1029-8479 (2013).

[42]W. Beenakker, S. Dittmaier, M. Krämer, B. Plümper, M. Spira, and P. Zerwas, NLO QCD corrections to $t\bar{t}H$ production in hadron collisions, Nuclear Physics B **653**, 151 –203, ISSN: 0550-3213 (2003).

[43]F. Bezrukov and M. Shaposhnikov, Why should we care about the top quark Yukawa coupling?, Journal of Experimental and Theoretical Physics **120**, 335–343, ISSN: 1090-6509 (2015).

[44]G. Watt and R. S. Thorne, Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs, Journal of High Energy Physics **2012**, 52, ISSN: 1029-8479 (2012).

[45]G. Aad, B. Abbott, et al., Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, Journal of High Energy Physics **2016**, 160, ISSN: 1029-8479 (2016).

[46]G. Aad, B. Abbott, et al., Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector, Physics Letters B **749**, 519 –541, ISSN: 0370-2693 (2015).

[47]Combination of the searches for Higgs boson production in association with top quarks in the $\gamma\gamma$, multilepton, and $b\bar{b}$ decay channels at $\sqrt{s}$=13 TeV with the ATLAS Detector, tech. rep. ATLAS-CONF-2016-068 (CERN, Geneva, 2016).

[48]Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector, tech. rep. ATLAS-CONF-2016-058 (CERN, Geneva, 2016).

[49]Search for ttH production in multilepton final states at $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-HIG-15-008 (CERN, Geneva, 2016).

[50]V. Khachatryan, A. M. Sirunyan, A. Tumasyan, et al., Search for the associated production of the Higgs boson with a top-quark pair, Journal of High Energy Physics **2014**, 87, ISSN: 1029-8479 (2014).

[51]G. Aad, B. Abbott, et al., Search for produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector, Physics Letters B **740**, 222 –242, ISSN: 0370-2693 (2015).

[52]Measurement of fiducial, differential and production cross sections in the $H \to \gamma\gamma$ decay channel with 13.3 fb^{-1} of 13 TeV proton-proton collision data with the ATLAS detector, tech. rep. ATLAS-CONF-2016-067 (CERN, Geneva, 2016).

[53]Updated measurements of Higgs boson production in the diphoton decay channel at $\sqrt{s} =$ 13 TeV in pp collisions at CMS. Tech. rep. CMS-PAS-HIG-16-020 (CERN, Geneva, 2016).

[54]Search for t$\bar{\text{t}}$H production in the H $\to$ b$\bar{\text{b}}$ decay channel with 2016 pp collision data at $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-HIG-16-038 (CERN, Geneva, 2016).

[55]G. Aad, B. Abbott, J. Abdallah, et al., Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, The European Physical Journal C **75**, 349, ISSN: 1434-6052 (2015).

[56]Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, tech. rep. ATLAS-CONF-2016-080 (CERN, Geneva, 2016).

[57]V. Khachatryan, A. M. Sirunyan, A. Tumasyan, et al., Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method, The European Physical Journal C **75**, 251, ISSN: 1434-6052 (2015).

[58]Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment, tech. rep. CMS-PAS-HIG-16-004 (CERN, Geneva, 2016).

[59]Search for associated production of Higgs bosons and top quarks in multilepton final states at $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-HIG-16-022 (CERN, Geneva, 2016).

[60]Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-HIG-17-004 (CERN, Geneva, 2017).

[61]L. Evans and P. Bryant, LHC Machine, Journal of Instrumentation **3**, S08001 (2008).

[62]CMS Collaboration, The CMS experiment at the CERN LHC, Journal of Instrumentation **3**, S08004 (2008).

[63]CMS Collaboration, CMS Luminosity – Public Results, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults, [Online; retrieved 16-April-2017], 2017.

[64]W. Herr and B Muratori, Concept of luminosity, (2006).

[65]Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-FSQ-15-005 (CERN, Geneva, 2016).

[66]CMS Collaboration, CMS Detector Drawings, CMS Document 5955-v1 [Online; retrieved 18-April-2017], 2012.

[67]A. Rao and S. Pandolfi, CMS pixel tracker transplant: everything went well so far, (2017).

[68]A. Dominguez, D. Abbaneo, K. Arndt, et al., CMS Technical Design Report for the Pixel Detector Upgrade, tech. rep. CERN-LHCC-2012-016. CMS-TDR-11 (2012).

[69]V. Khachatryan, A. M. Sirunyan, A. Tumasyan, et al., Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV, Journal of Instrumentation **10**, P06005 (2015).

[70]F. Thyssen, Performance of the Resistive Plate Chambers in the CMS experiment, Journal of Instrumentation **7**, C01104 (2012).

[71]CMS Collaboration, CMS SoftWare, https://github.com/cms-sw/cmssw, [Online, retrieved 24-April-2017], 2017.

[72]M. Jeitler, Upgrade of the trigger system of CMS, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **718**, 11–15, ISSN: 0168-9002 (2013).

[73]S. Cittolin, A. Rácz, and P. Sphicas, CMS The TriDAS Project: Technical Design Report, Vol. 2: Data Acquisition and High-Level Trigger, tech. rep. (Geneva, 2002).

[74]R. Brun and F. Rademakers, ROOT: An object oriented data analysis framework, Nuclear Instruments and Methods in Physics Research **A389**, 81–86 (1997).

[75]A. Hoecker, P. Speckmayer, J. Stelzer, et al., TMVA - Toolkit for Multivariate Data Analysis, ArXiv Physics e-prints (2007).

[76]F. Pedregosa, G. Varoquaux, A. Gramfort, et al., Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research **12**, 2825–2830, ISSN: 1532-4435 (2011).

[77]CMS Collaboration, CMS DQM service, https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions16/13TeV/ReReco/Final/Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt, [Online, retrieved 26-April-2017], 2016.

[78]L. Tuura, G. Eulisse, and A. Meyer, CMS data quality monitoring web service, Journal of Physics: Conference Series **219**, 072055 (2010).

[79]V. Kuznetsov, D. Evans, and S. Metson, The CMS Data Aggregation System, Procedia Computer Science **1**, 1535 –1543, ISSN: 1877-0509 (2010).

[80]J. Alwall, R. Frederix, S. Frixione, et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, Journal of High Energy Physics **2014**, 79, ISSN: 1029-8479 (2014).

[81]S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, Journal of High Energy Physics **2010**, 43, ISSN: 1029-8479 (2010).

[82]T. Sjöstrand, S. Ask, J. R. Christiansen, et al., An introduction to PYTHIA 8.2, Computer Physics Communications **191**, 159 –177, ISSN: 0010-4655 (2015).

[83]S. Agostinelli, J. Allison, K. Amako, et al., GEANT4–a simulation toolkit, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506**, 250–303, ISSN: 0168-9002 (2003).

[84]J. Allison, K. Amako, J. Apostolakis, and H. others, Geant4 Developments and Applications, IEEE Transactions on Nuclear Science **53**, 270–278, ISSN: 0018-9499 (2006).

[85]J. Allison, K. Amako, J. Apostolakis, et al., Recent developments in GEANT4, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **835**, 186–225, ISSN: 0168-9002 (2016).

[86]CMS VHbb team, CMSSW SoftWare, https://github.com/vhbb/cmssw/, [Online, retrieved 01-May-2017], 2017.

[87]Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector, tech. rep. CMS-PAS-PFT-10-001 (2010).

[88]CMS Collaboration, Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV, Journal of Instrumentation **10**, P06005 (2015).

[89]W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC, Journal of Physics G: Nuclear and Particle Physics **31**, N9 (2005).

[90]EGamma Physics Object Group, Multivariate Electron Identification for Run2, https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentificationRun2, [Online (restricted access), retrieved 27-April-2017], 2017.

[91]The CMS collaboration, Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV, Journal of Instrumentation **7**, P10002 (2012).

[92]CMS Muon POG, Baseline muon selections for Run-II, https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2, [Online (restricted access), retrieved 30-April-2017], 2017.

[93] TRK report – HIP Mitigation, CMS Week 20-24 June 2016 [Online (restricted access), retrieved 01-May-2017], 2016.

[94] M. Cacciari, G. P. Salam, and G. Soyez, The anti-k_t jet clustering algorithm, Journal of High Energy Physics **2008**, 063 (2008).

[95] The CMS collaboration, Determination of jet energy calibration and transverse momentum resolution in CMS, Journal of Instrumentation **6**, P11002 (2011).

[96] CMS JERC subgroup, Recommended Jet Energy Corrections and Uncertainties For Data and MC, https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC, [Online (restricted access), retrieved 01-May-2017], 2017.

[97] CMS BTV POG, Usage of b/c Tag Objects for 13 TeV Data in 2016 and 80X MC, https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation80XReReco, [Online (restricted access), retrieved 29-April-2017], 2017.

[98] The CMS collaboration, Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS, Journal of Instrumentation **11**, P01019 (2016).

[99] Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2, tech. rep. CMS-PAS-TAU-16-002 (CERN, Geneva, 2016).

[100] A. Nayak, A. Raspereza, J. Steggemann, Y. Takahashi, and C. Veelken, Multivariate τ-identification algorithms for 13 TeV data, tech. rep. CMS Analysis Note 2015/310, [Online (restricted access); retrieved 03-May-2017] (CERN, Geneva, 2015).

[101] Measurement of the $t\bar{t}H$ cross section in the $H \to \tau\tau$ decay channel at $\sqrt{s} = 13\,\text{TeV}$ with the matrix element method, tech. rep. CMS AN-2016/273 v10, [Online (restricted access); retrieved 12-May-2017] (CERN, Geneva, 2017).

[102] Search for the standard model Higgs boson produced in association with top quarks in multilepton final states, tech. rep. CMS AN-2013/159 v12, [Online (restricted access); retrieved 07-May-2017] (CERN, Geneva, 2013).

[103] N. Adam, J. Berryhill, V. Halyo, A. Hunt, and M. K., Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data, tech. rep. CMS AN-2009/111, [Online (restricted access); retrieved 03-May-2017] (CERN, Geneva, 2009).

[104] CMS Tau POG, TauID for 13 TeV run: recommendation from the Tau POG, https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendation13TeV, [Online, retrieved 03-May-2017], 2017.

[105] CMS BTV POG, Event reweighting using scale factors calculated with a tag and probe method, https://twiki.cern.ch/twiki/bin/view/CMS/BTagShapeCalibration, [Online, retrieved 03-May-2017], 2017.

[106] Luminosity Physics Object Group, Current recommendations for luminosity estimations – 2016 pp run, https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM, [Online, retrieved 08-May-2017], 2017.

[107] LHC Higgs Cross Section Working Group, ttH/tHq Higgs Associated Production Process, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGTTH, [Online, retrieved 08-May-2017], 2016.

[108] S. Dawson, L. H. Orr, L. Reina, and D. Wackeroth, Next-to-leading order QCD corrections to $pp \to t\bar{t}h$ at the CERN Large Hadron Collider, Phys. Rev. D **67**, 071503 (2003).

[109] K. Kondo, Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. I. Method and Toy Models, Journal of the Physical Society of Japan **57**, 4126–4140 (1988).

[110]K. Kondo, Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. II. Mass Spectra for $2 \to 2$ Processes, Journal of the Physical Society of Japan **60**, 836–844 (1991).

[111]Search for the associated $t\bar{t}H(\to b\bar{b})$ production using the Matrix Element Method, tech. rep. CMS AN-2013/313 v7, [Online (restricted access); retrieved 24-May-2017] (CERN, Geneva, 2017).

[112]L. Bianchini, B. Calpas, J. Conway, A. Fowlie, L. Marzola, L. Perrini, and C. Veelken, Reconstruction of the higgs mass in events with higgs bosons decaying into a pair of leptons using matrix element techniques, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **862**, 54 –84, ISSN: 0168-9002 (2017).

[113]R. Neal, Probabilistic Inference using Markov Chain Monte Carlo Methods, tech. rep. CRG-TR-93-1, [Online, retrieved 24-May-2017] (Department of Computer Science, University of Toronto, 1993).

[114]R. Neal, Bayesian Training of Backpropagation Networks by the Hybrid Monte Carlo Method, tech. rep. CRG-TR-92-1, [Online, retrieved 24-May-2017] (Department of Computer Science, University of Toronto, 1992).

[115]G. P. Lepage, A new algorithm for adaptive multidimensional integration, Journal of Computational Physics **27**, 192 –203, ISSN: 0021-9991 (1978).

[116]T. Ohl, Vegas revisited: Adaptive Monte Carlo integration beyond factorization, Computer Physics Communications **120**, 13 –19, ISSN: 0010-4655 (1999).

[117]The CMS collaboration, Missing transverse energy performance of the CMS detector, Journal of Instrumentation **6**, P09001 (2011).

[118]M. Cacciari and G. P. Salam, Pileup subtraction using jet areas, Physics Letters B **659**, 119–126, ISSN: 0370-2693 (2008).

[119]M. Cacciari, G. P. Salam, and G. Soyez, The catchment area of jets, Journal of High Energy Physics **2008**, 005 (2008).

[120]CMS Collaboration, Description and performance of track and primary-vertex reconstruction with the CMS tracker, Journal of Instrumentation **9**, P10009 (2014).

[121]CMS Collaboration, Identification of b-quark jets with the CMS experiment, Journal of Instrumentation **8**, P04013 (2013).

[122]Procedure for the LHC Higgs boson search combination in Summer 2011, tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11 (CERN, Geneva, 2011).

[123]L. Moneta, K. Belasco, K. Cranmer, et al., Proceedings of Science **ACAT2010**, 057 (2010).

[124]CMS Collaboration, HiggsAnalysis-CombinedLimit, https://github.com/cms-analysis/HiggsAnalysis-CombinedLimit, [Online, retrieved 11-May-2017], 2017.

[125]CMS Collaboration, CombineHarvester, https://github.com/cms-analysis/CombineHarvester, [Online, retrieved 11-May-2017], 2017.

[126]LHC Higgs Cross Section Working Group, SM Higgs cross section and associated uncertainties, https://twiki.cern.ch/twiki/pub/LHCPhysics/LHCHXSWG/Higgs_XSBR_YR4_update.xlsx, [Online; retrieved 13-April-2017], 2016.

Abbreviations

AOD	Analysis Object Data
APD	avalanche photodiode
ATLAS	A Toroidal LHC ApparatuS
BDT	boosted decision tree
BPIX	barrel pixel detector
BSM	beyond the SM
BTV	*b*-tagging and vertex fitting
c.f.	check for
CKM	Cabbibo-Kobayashi-Maskawa
CL	confidence level
CMS	Compact Muon Solenoid
c.o.m	center-of-mass
CSC	cathode strip chamber
CSV	Combined Secondary Vertex
DAQ	data acquisition
DAS	Data Aggregation System
d.o.f	degrees of freedom
DQM	Data Quality Monitoring
DT	drift tube
EB	ECAL Barrel
ECAL	electromagnetic calorimeter
EE	ECAL Endcap
EW	electroweak
FF	fake factor
FPIX	forward pixel detector
FR	fake rate
FWHM	full width at half maximum
GGF	gluon-gluon fusion
GSF	Gaussian Sum Filter
HB	HCAL barrel
h.c.	hermitian conjugate
HCAL	hadronic calorimeter
HE	HCAL endcap
HEP	high energy physics
HF	HCAL forward
HLT	high-level trigger
HPS	hadrons plus strips
HPD	hybrid photodiode
IP	impact parameter
IR	interaction region
ISR	initial state radiation
JEC	jet energy correction
JERC	jet energy resolution and corrections
LEP	Large Electron-Positron Collider
LHC	Large Hadron Collider
LIPS	Lorentz-invariant phase space
LO	leading order
MB	muon barrel
MC	Monte Carlo
ME	muon endcap
MEM	matrix element method
MET	missing transverse energy
MHT	missing $\vec{H}_T$, $\vec{\not{H}}_T$
ML	maximum likelihood
MVA	multivariate analysis
NLO	next-to-LO
NNLO	next-to-NLO
N3LO	next-to-NNLO
1PI	one-particle irreducible
PD	primary dataset
p.d.f	probability distribution function
PDF	parton distribution function
PF	Particle Flow
PMT	photomultiplier
POG	Physics Object Group
POI	parameter of interest
PS	Proton Synchroton
PSB	Proton Synchroton Booster
PU	pileup
PV	primary vertex
QCD	quantum chromodynamics
ROC	readout chip
RPC	resistive plate chamber
SFOS	same-flavor opposite sign
SIP	significance of IP
SL	superlayer
SM	Standard Model
SPS	Super Proton Synchroton
$\sqrt{s}$	c.o.m energy
SR	signal region
SS	same-sign
SSB	spontaneous symmetry breaking
SV	secondary vertex
TEC	Tracker EndCap
TF	transfer function
TIB	Tracker Inner Barrel

TID	Tracker Inner Disk	**VEV**	vacuum expectation value
TOB	Tracker Outer Barrel	**VPT**	vacuum phototriode
UV	ultraviolet	**WP**	working point
VBF	vector boson fusion	**w.r.t**	with respect to

Acknowledgments

First, I would like to thank all of my colleagues at NICPB, especially my supervisor Christian Veelken, who was always available for discussion and provided detailed answers to my questions. I thank Mario Kadastik for giving me an opportunity to participate in the particle physics analysis. I also thank Luca Marzola who was again willing to take up the task of co-supervising me.

Big up for my family, friends, dorm mates and peers who were willing to listen about my experiences in the field. Also, cheers to theorists at UT and members of the CMS collaboration.

Aitäh!

A. The Standard Model

A.1. The Higgs doublet

The assignment $Y = 1$ is motivated by the Gell-Mann-Nishijima formula

$$Q = T_3 + \frac{Y}{2}\,, \tag{A.1}$$

which relates electric charge Q to weak isospin T_3 and hypercharge Y. The weak isospin refers to the eigenvalue of the diagonal (commonly third) $SU(2)_L$ generator, which equals to $T_3 = \pm\frac{1}{2}$ for the top and bottom component of the Higgs $SU(2)_L$ doublet (2.20)

$$\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}\,, \tag{A.2}$$

respectively. Since we require that the VEV of the Higgs field and the Higgs boson have $Q = 0$, we set $Y = 1$. Consequently, a positive electric charge ($Q = +1$) is assigned to the field ϕ^+ because its weak isospin is equal to $T_3 = +\frac{1}{2}$.

There also exists Higgs doublet

$$\bar{\phi} \equiv \mathcal{C}\phi = i\sigma_2\phi^* = \begin{pmatrix} \phi^{0*} \\ -\phi^{+*} \end{pmatrix} \overset{!}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} v+h \\ 0 \end{pmatrix}\,, \tag{A.3}$$

which transforms the same way as the charge-conjugate of ϕ:

$$\begin{aligned} \bar{\phi} \to \bar{\phi}' = e^{-ig_2\mathbf{W}\boldsymbol{\sigma}/2}\bar{\phi} = e^{-ig_2\mathbf{W}\boldsymbol{\sigma}/2} i\sigma_2\phi^* = i\sigma_2 e^{ig_2\mathbf{W}\boldsymbol{\sigma}^*/2}\phi^* = i\sigma_2\left(e^{-ig_2\mathbf{W}\boldsymbol{\sigma}/2}\phi\right)^* = \\ = i\sigma_2\left(\phi'\right)^*\,, \end{aligned} \tag{A.4}$$

where we used the identity

$$e^{ig_2\mathbf{W}\boldsymbol{\sigma}^*/2} = \sigma_2 e^{-ig_2\mathbf{W}\boldsymbol{\sigma}/2}\sigma_2^\dagger \tag{A.5}$$

based on the fact that $\sigma_2\sigma_i^* = -\sigma_i\sigma_2$. The charge-conjugation operator $\mathcal{C}$ inverts the signs of the weak isospins, i.e. $T_3(\phi^{+*}) = -T_3(\phi^+)$ and $T_3(\phi^{0*}) = -T_3(\phi^0)$. The sign of the hypercharge is also reversed, $Y = -1$, which in turn means that ϕ^{+*} obtains negative electric charge, but the other field ϕ^{0*} remains neutral.

A.2. Gauge-kinetic term of the Higgs field

Let us expand the gauge kinetic term (2.23) in more detailed fashion:

$$
\begin{aligned}
|D\phi|^2 &= \frac{1}{2}\left|\left(\partial - ig_2\mathbf{W}\cdot\mathbf{T} - \frac{ig_1}{2}B\right)\begin{pmatrix}0\\ v+h\end{pmatrix}\right|^2 = \qquad (A.6)\\
&= \left[\mathbf{T} = \frac{1}{2}\boldsymbol{\sigma} = \frac{1}{2}\left(\begin{pmatrix}0&1\\1&0\end{pmatrix},\ \begin{pmatrix}0&-i\\i&0\end{pmatrix},\ \begin{pmatrix}1&0\\0&-1\end{pmatrix}\right)\right] = \\
&= \frac{1}{2}\left|\begin{pmatrix}\partial - \frac{ig_2}{2}W_3 - \frac{ig_1}{2}B & -\frac{ig_2}{2}(W_1 - iW_2)\\ -\frac{ig_2}{2}(W_1+iW_2) & \partial + \frac{ig_2}{2}W_3 - \frac{ig_1}{2}B\end{pmatrix}\begin{pmatrix}0\\v+h\end{pmatrix}\right|^2 = \\
&= \frac{1}{2}\left|\begin{matrix}-\frac{ig_2}{2}(W_1 - iW_2)(v+h)\\ \partial h + \frac{i}{2}(g_2W_3 - g_1B)(v+h)\end{matrix}\right|^2 = \\
&= \frac{1}{2}(\partial h)^2 + \frac{g_2^2}{8}(W_1+iW_2)(W_1-iW_2)(v+h)^2 + \frac{1}{8}(g_2W_3 - g_1B)^2(v+h)^2 = \\
&= \frac{1}{2}(\partial h)^2 + \frac{g_2^2}{4}W^-W^+(v+h)^2 + \frac{g_1^2+g_2^2}{8}Z^2(v+h)^2\,.
\end{aligned}
$$

To see why the identification (2.24) for $W^\pm$ bosons works, let us write out relevant part of the covariant derivative (2.10):

$$
\begin{aligned}
W_1T_1 + W_2T_2 &= \frac{W_1 - iW_2}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}0&1\\0&0\end{pmatrix} + \frac{W_1+iW_2}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}0&0\\1&0\end{pmatrix} = \\
&= W^+T^+ + W^-T^-\,, \qquad (A.7)
\end{aligned}
$$

where

$$
T_+ = \frac{1}{\sqrt{2}}\sigma_+ = \frac{1}{\sqrt{2}}\begin{pmatrix}0&1\\0&0\end{pmatrix} \quad \text{and} \quad T_- = \frac{1}{\sqrt{2}}\sigma_- = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0\\1&0\end{pmatrix} \qquad (A.8)
$$

are the ladder operators (or non-Cartan generators) of the $SU(2)_L$ group, the definition of which ensures the structure of Cartan subalgebra, $[T_+, T_-] = T_3$ (hence the normalization factors in the definition of $W^\pm$, (2.24)). Let us consider interaction term between left-handed quarks in (2.1) with the expansion (A.7):

$$
\begin{aligned}
i\bar{Q}_L\not{D}Q_L \supset g_2\bar{Q}_L(\not{W}_1T_1 + \not{W}_2T_2)Q_L = \qquad (A.9)\\
= \frac{g_2}{2}\left[\begin{pmatrix}\bar{u}^i & \bar{d}^i\end{pmatrix}_L(\not{W}_1 - i\not{W}_2)\sigma_+\begin{pmatrix}u^i\\d^i\end{pmatrix}_L + \begin{pmatrix}\bar{u}^i & \bar{d}^i\end{pmatrix}_L(\not{W}_1 + i\not{W}_2)\sigma_-\begin{pmatrix}u^i\\d^i\end{pmatrix}_L\right] = \\
= \frac{g_2}{2}\left[\bar{u}^i_L(\not{W}_1 - i\not{W}_2)d^i_L + \bar{d}^i_L(\not{W}_1 + i\not{W}_2)u^i_L\right] = \\
= \frac{g_2}{\sqrt{2}}\left(\bar{u}^i_L\not{W}^+d^i_L + \bar{d}^i_L\not{W}^-u^i_L\right)\,.
\end{aligned}
$$

Given that the u- and d-quarks have charges $Q = +\frac{2}{3}$ and $Q = -\frac{1}{3}$, the conservation of electric charge in weak interactions implies neutrally charged interaction terms, which leads us to the anticipated prescription for $W^\pm$ where the superscript also denotes the electric charge of the boson.

Although the definition of Z in (2.24) is motivated by the last term of (2.23), the normalization constant of the new field needs some explaining. Let us rewrite the definitions of A and Z in terms of Weinberg angle θ_w,

$$\begin{pmatrix} A \\ Z \end{pmatrix} \equiv R(\theta_w) \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad \text{with} \quad R(\theta_w) = \begin{pmatrix} c_w & s_w \\ -s_w & c_w \end{pmatrix}, \tag{A.10}$$

where

$$s_w = \sin\theta_w \equiv \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \quad \text{and} \quad c_w = \cos\theta_w \equiv \frac{g_2}{\sqrt{g_1^2 + g_2^2}}. \tag{A.11}$$

In principle, the couplings can be visualized as the sides of a right triangle with angle θ_w between g_2 and hypotenuse, as shown in fig. A.1.

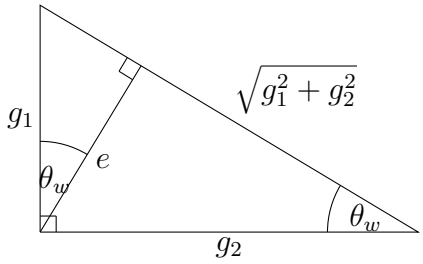


Figure A.1: Visualization of electric charge e and $SU(2)_L \times U(1)_Y$ couplings $g_2 \approx 0.65$ and $g_1 \approx 0.35$ in relation to the Weinberg angle $\theta_w \approx 28.7°$ [24]. All coupling depicted here are dimensionless.

In the basis of fields B and W_3, the new fields are orthogonal by construction:

$$Z \cdot A = \begin{pmatrix} c_w & -s_w \end{pmatrix} \begin{pmatrix} s_w \\ c_w \end{pmatrix} = c_w s_w - s_w c_w = 0\,. \tag{A.12}$$

The above prescription allows to express gauge fields in their mass eigenstates by multiplying (A.10) on the left by $R(-\theta_w)$, which in turn gives the following EW part of the covariant derivative (2.10):

$$D \supset -ig_2 W_3 T_3 - ig_1 B\frac{Y}{2} = -i\left(g_2 s_w T_3 + g_1 c_w \frac{Y}{2}\right) A - i\left(g_2 c_w T_3 - g_1 s_w \frac{Y}{2}\right) Z\,. \tag{A.13}$$

We can rearrange the factor in front of the photon field A:

$$g_2 s_w T_3 + g_1 c_w \frac{Y}{2} = \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}}\left(T_3 + \frac{Y}{2}\right) \overset{(A.1)}{=} \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}} Q \equiv eQ\,, \tag{A.14}$$

where we have identified electric charge e and charge operator Q. Therefore, the full covariant derivative in the mass eigenstates of the gauge bosons reads

$$D = -ig_S \mathbf{Gt} - \frac{ie}{s_w}(W^+ T^+ + W^- T^-) - \frac{ie}{s_w c_w}(T_3 - s_w^2 Q) Z - ieQA\,. \tag{A.15}$$

The SSB has now become apparent: $SU(3)_C \times SU(2)_L \times U(1)_Y \overset{\text{SSB}}{\to} SU(3)_C \times U(1)_{\text{EM}}$, i.e. from the EW sector only the electromagnetic gauge group $U(1)_{\text{EM}}$ emerges after the SSB.

A.3. Potential term of the Higgs field

Here we try to reproduce the result (2.26) by identifying $\mu^2 = -\lambda v^2$ from (2.17):

$$\begin{aligned} V(\phi) &= -\frac{\lambda}{2}v^2(v+h)^2 + \frac{\lambda}{4}(v+h)^4 = \\ &= -\frac{\lambda}{2}v^2(h^2+2hv+v^2) + \frac{\lambda}{4}(h^4+4h^3v+6h^2v^2+4hv^3+v^4) = \\ &= \frac{\lambda}{4}h^4 + \lambda h^3 v + \lambda h^2 v^2 \,. \end{aligned} \tag{A.16}$$

In the last line we omitted terms containing only the VEV since $v = const$ and the Lagrangian density is invariant under shifts by constant quantities.

A.4. Quark sector

Let us dissect the Yukawa contribution between the quarks and Higgs in more detail by explicitly expanding the terms of (2.31):

$$\begin{aligned} \mathcal{L}_{q,\text{Yukawa}} &= -y^u_{ij}\bar{u}^i_R\bar{\phi}^\dagger Q^j_L - y^d_{ij}\bar{d}^i_R\phi^\dagger Q^j_L + \text{h.c.} = \\ &= -\frac{v+h}{\sqrt{2}}\left(y^u_{ij}\bar{u}^i_R u^j_L + y^d_{ij}\bar{d}^i_R d^j_L + y^{u*}_{ji}\bar{u}^j_L u^i_R + y^{d*}_{ji}\bar{d}^j_L d^i_R\right) . \end{aligned} \tag{A.17}$$

By using the relations (2.32)-(2.34), we can obtain the mass terms:

$$\begin{aligned} -\mathcal{L}_{q,m} &= \bar{u}^{k'}_R u^{l'}_L \mathcal{M}^u_{ij} U^{R*}_{ki} U^L_{jl} + \bar{u}^{k'}_L u^{l'}_R \mathcal{M}^{u*}_{ji} U^{L*}_{kj} U^R_{il} + \bar{d}^{k'}_R d^{l'}_L \mathcal{M}^d_{ij} D^{R*}_{ki} D^L_{jl} + \bar{d}^{k'}_L d^{l'}_R \mathcal{M}^{d*}_{ji} D^{L*}_{kj} D^R_{ik} = \\ &= \bar{u}^{k'}_R u^{l'}_L (U^\dagger_R \mathcal{M}^u U_L)_{kl} + \bar{d}^{k'}_R d^{l'}_L (D^\dagger_R \mathcal{M}^d D_L)_{kl} + \text{h.c.} = \\ &= \hat{\mathcal{M}}^u(\bar{u}'_R u_L + \bar{u}'_L u_R) + \hat{\mathcal{M}}^d(\bar{d}'_R d_L + \bar{d}'_L d_R) = \\ &= \hat{\mathcal{M}}^u \bar{u}' u' + \hat{\mathcal{M}}^d \bar{d}' d' \equiv \\ &\equiv m_{qi}\bar{q}'_i q'_i \,. \end{aligned} \tag{A.18}$$

Higgs-to-quark coupling terms can be derived in similar fashion.

The unitary matrices U_L and U_R (as well as D_L and D_R) are independent from each other, and are chosen such that the biunitary transformations of $\mathcal{M}^u$ and $\mathcal{M}^d$ produce diagonal mass matrices with positive elements on their diagonals:

$$\begin{aligned} \hat{\mathcal{M}}^u &\equiv U^\dagger_R \mathcal{M}^u U_L = \text{diag}(m_u, m_c, m_t) \,, \\ \hat{\mathcal{M}}^d &\equiv D^\dagger_R \mathcal{M}^d D_L = \text{diag}(m_d, m_s, m_b) \,. \end{aligned} \tag{A.19}$$

The mass eigenstates of the quarks mix due to covariant derivative (A.15) entering in the interaction term between left-handed $SU(2)_L$ doublet, as we already saw in (A.9). It is easy to see why CKM matrix $V \equiv U^\dagger_L D_L$ is defined as such:

$$\begin{aligned} i\bar{Q}_L \not{D} Q_L \supset \frac{g_2}{\sqrt{2}}(\bar{u}^{k'}_L U^{L*}_{ki} D^L_{il} \not{W}^+ d^{l'}_L + \not{d}^{k'}_L D^{L*}_{ki} U^L_{il} \not{W}^- u^{l'}_L) = \\ = \frac{g_2}{\sqrt{2}}\left[\bar{u}^{k'}_L (U^\dagger_L D_L)_{kl} \not{W}^+ d^{l'}_L + \not{d}^{k'}_L (D^\dagger_L U_L)_{kl} \not{W}^- u^{l'}_L\right] = \\ = \frac{g_2}{\sqrt{2}}(\bar{u}'_L V \not{W}^+ d'_L + \bar{d}'_L V^\dagger \not{W}^- u'_L) \,. \end{aligned} \tag{A.20}$$

The magnitude of CKM elements are [24, p 229]

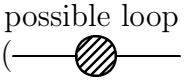
$$V = \begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97434 & 0.22506 & 0.00357 \\ 0.22492 & 0.97351 & 0.0411 \\ 0.00875 & 0.0403 & 0.99915 \end{bmatrix}, \tag{A.21}$$

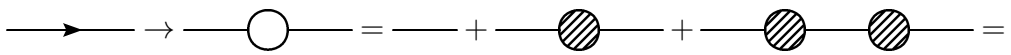
where the probability for flavor transition $i \leftrightarrow j$ is given by $|V_{ij}|^2$. For instance, the probability for a top quark not decaying into a bottom quark is less than 0.2%. The neutral current contribution involves operator $T_3 - s_w^2 Q$, which is diagonal and hence does not mix the mass eigenstates.

A.5. Decay widths and branching ratios

Bare propagator of an unstable particle A,

$$= \frac{i}{p^2 - m^2}\,, \tag{A.22}$$

could introduce unphysical singularities to probability amplitudes $\mathcal{M}(X \to A \to Y)$ (and thereby to observables like cross sections) if the particle goes to mass shell ($p^2 \to m^2$). It turns out that the propagator (A.22) must be modified to include all higher order contributions that suppress these singularities. This can be done by resumming all possible loop diagrams which we encapsulate into a sum of one-particle irreducible (1PI) ( $\equiv -i\Sigma(p^2)$)[1] corrections to the bare propagator:

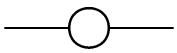


$$= \frac{i}{p^2 - m_R^2} \sum_{n=0}^{\infty} \left[\frac{\Sigma(p^2)}{p^2 - m_R^2} \right]^n = \frac{i}{p^2 - m_R^2 - \Sigma(p^2)}\,, \tag{A.23}$$

where m_R refers to renormalized mass subject to on-shell renormalization conditions $\Sigma(m_R^2) = \Sigma'(m_R^2) = 0$. For an unstable particle the sum of 1PI diagrams may acquire complex component, which shifts the mass pole to complex plane if we set the denominator of (A.23) to zero. Thus, it is custom to define real mass pole (or Breit-Wigner mass) m_P as the solution to

$$m_R^2 - m_P^2 - \mathrm{Re}\Sigma(m_P^2) = 0\,, \tag{A.24}$$

which brings us to the following expression for the full propagator:



$$= \frac{i}{p^2 - m_P^2 - i\mathrm{Im}\Sigma(p^2)}\,. \tag{A.25}$$

We define the (partial) decay width $\Gamma(A \to X)$ for the process $A \to X$ as

$$\Gamma(A \to X) \equiv \frac{1}{2m_A} \int \mathrm{dLIPS}_X \, |\mathcal{M}(A \to X)|^2 \tag{A.26}$$

[1] One cannot split 1PI diagram into two parts without crossing more than one propagator.

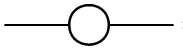
and the total decay width of A as a sum of decay widths for all possible processes,

$$\Gamma_A \equiv \Gamma_{\text{tot}}(A \to \text{any}) = \sum_X \Gamma(A \to X)\,. \tag{A.27}$$

According to optical theorem, the imaginary part of the forward scattering amplitude is related to the total decay width as follows:

$$2\text{Im}\mathcal{M}(A \to A) = \sum_X \int \text{dLIPS}_X\, |\mathcal{M}(A \to X)|^2 = \frac{\Gamma_{\text{tot}}}{m_A}\,. \tag{A.28}$$

Therefore, by identifying $\mathcal{M}(A \to A) = \Sigma(p^2)$, the propagator (A.25) takes the form



$$= \frac{i}{p^2 - m_P^2 + i m_P \Gamma_{\text{tot}}} \tag{A.29}$$

also known as Breit-Wigner propagator. The modified propagator enters the cross sections in quadrature,

$$\sigma \sim \left| \frac{i}{p^2 - m_P^2 + i m_P \Gamma_{\text{tot}}} \right|^2 = \frac{1}{(p^2 - m_P^2)^2 + (m_P \Gamma_{\text{tot}})^2}\,, \tag{A.30}$$

thus following so-called Breit-Wigner (or Cauchy-Lorentz) distribution. In other words, the unstable particle manifests itself as a resonance in the cross section, the distribution of which peaks at mass m_P and has full width at half maximum (FWHM) of $2m_P\Gamma_{\text{tot}}$. If $\Gamma_{\text{tot}} \ll m_P$ we obtain the δ-function approximation of the Breit-Wigner propagator:

$$\frac{1}{(p^2 - m_P^2)^2 + (m_P \Gamma_{\text{tot}})^2} \overset{\Gamma_{\text{tot}} \ll m_P}{\longrightarrow} \frac{\pi}{\Gamma_{\text{tot}} m_P} \delta^2(p^2 - m_P^2)\,. \tag{A.31}$$

On the other hand, larger decay width means smaller mean lifetime τ of the particle:

$$\tau = \frac{1}{\Gamma_{\text{tot}}}\,. \tag{A.32}$$

The result (A.29) is valid not only for a scalar field but spinor and vector field as well.

The decay widths help us to define branching ratio $\text{Br}(A \to X)$ as a ratio of its decay width to the total one:

$$\text{Br}(A \to X) \equiv \frac{\Gamma(A \to X)}{\Gamma_{\text{tot}}}\,, \tag{A.33}$$

which ensures that $\sum_X \text{Br}(A \to X) = 1$. The branching ratios can be interpreted as the probability for process $A \to X$ occurring relative to all possible processes ($A \to$ any).

B. CMS coordinate system and kinematics

B.1. Coordinate system

The Cartesian coordinate system used in CMS is right-handed: x-axis points towards the center of the LHC ring, z-axis directed along Beam 2 (i.e. the beam moving anti-clockwise, see fig. 3.1) and y-axis is perpendicular to the x-z plane, pointing upwards, with the origin of all axes set to the nominal interaction point in the center of the detector. The fact that the detector is radially symmetric around the z-axis (i.e. around the beam line) motivates the use of another coordinate system[1] with the following definitions:

$$r \equiv \sqrt{x^2 + y^2}\,, \quad \theta \equiv \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}}\,, \quad \phi \equiv \arctan \frac{y}{x}\,, \tag{B.1}$$

where the radial distance $r \in [0, \infty)$ is measured from the z-axis, the polar angle $\theta \in [0, \pi)$ from the positive z direction (i.e. it denotes the opening angle to the z axis) and the azimuthal angle $\phi \in (-\pi, \pi]$ indicates the orientation in the y-x plane from positive x direction. The relationship between Cartesian and these new coordinates is visualized in fig. B.1(a).

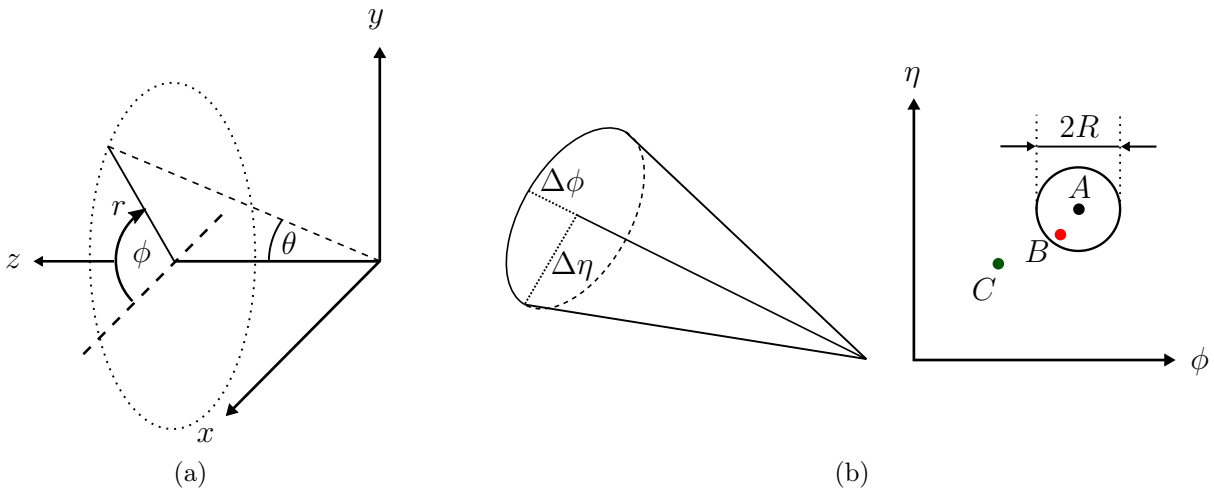


Figure B.1: (a) the CMS coordinate systems: Cartesian (x, y, z) and alternative (r, θ, ϕ) coordinates; (b) isolation cone centered on momentum vector of particle (left) and its size in angular space (right) (see discussion of (B.13)).

The pseudorapidity η for convenience, defined by

$$\eta \equiv -\ln \tan \frac{\theta}{2}\,, \quad \text{or conversely} \quad \theta = 2 \arctan e^{-\eta}\,, \tag{B.2}$$

is then used instead of the polar angle θ. The advantage of pseudorapidity is that it is invariant under Lorentz boosts in z direction, i.e. the difference in pseudorapidity ($\Delta\eta = \eta_1 - \eta_2$) does not depend on how much energy is carried away by (unmeasured) proton remnants.

[1] Oftentimes this coordinate system is referred to as „cylindrical coordinates" but that is not exactly what we use here.

It is easy to see that $\eta = 0$ if $\theta = \frac{\pi}{2}$ and approaches to $+\infty$ $(-\infty)$ if $\theta \to 0$ $(\theta \to \pi)$. Notice that

$$\sinh\eta = \frac{e^{\eta} - e^{-\eta}}{2} = \frac{1}{2}\left(\frac{1}{\tan\frac{\theta}{2}} - \tan\frac{\theta}{2}\right) = \cot\theta \overset{\text{(B.1)}}{=} \cot\arccos\frac{z}{\sqrt{r^2+z^2}} =$$

$$= \left.\frac{A}{\sqrt{1-A^2}}\right|_{A=z/\sqrt{r^2+z^2}} = \frac{z}{r} \quad \Rightarrow \quad z = r\sinh\eta\,, \tag{B.3}$$

$$\sin\phi \overset{\text{(B.1)}}{=} \sin\arctan\frac{y}{x} = \left.\frac{A}{\sqrt{1+A^2}}\right|_{A=y/x} = \frac{y}{r} \quad \Rightarrow \quad y = r\sin\phi\,, \tag{B.4}$$

$$\cos\phi \overset{\text{(B.1)}}{=} \cos\arctan\frac{y}{x} = \left.\frac{1}{\sqrt{1+A^2}}\right|_{A=y/x} = \frac{x}{r} \quad \Rightarrow \quad x = r\cos\phi\,, \tag{B.5}$$

$$\sqrt{x^2+y^2+z^2} \overset{\text{(B.3)-(B.5)}}{=} \sqrt{r^2\cos^2\phi + r^2\sin^2\phi + r^2\sinh^2\eta} = r\cosh\eta\,. \tag{B.6}$$

Since the CMS detector measures the 4-momenta of particles, $p \equiv (E,\, p_x,\, p_y,\, p_z)$, the above coordinate prescription applies to these 4-vectors as well, in which the „radial" coordinate r of the ordinary space is now called transverse momentum p_T in momentum space:

$$p_T \equiv |\mathbf{p}_T| \equiv \sqrt{p_x^2 + p_y^2} \quad (\text{with } \mathbf{p}_T = (p_x,\, p_y))\,. \tag{B.7}$$

The longitudinal momentum is naturally p_z since its direction coincides with the beam line. Following the relations (B.3)-(B.6), it is easy to see that the vector $\mathbf{p} = (p_x,\, p_y,\, p_z)$ in Cartesian coordinates can be written in terms of the components $(p_T,\, \eta,\, \phi)$ as follows:

$$p_x = p_T\cos\phi\,, \quad p_y = p_T\sin\phi\,, \quad p_z = p_T\sinh\eta\,, \quad |\mathbf{p}| = p_T\cosh\eta\,. \tag{B.8}$$

Sometimes it is beneficiary to distinguish a given set of objects by their p_T, where the object with the highest p_T is referred to as „leading", the one with the second highest as „sub-leading" and the third highest as „third" object.

One can also construct the so-called transverse mass m_T of a single particle from its transverse momentum p_T and its invariant mass m:

$$m_T^2 \equiv m^2 + p_T^2 = E^2 - p_z^2\,. \tag{B.9}$$

Since quantities that are transverse to the direction of a Lorentz boost remain invariant under the boost, the transverse mass does not change under Lorentz boosts in beam direction.

B.2. Missing transverse energy

If we sum all momenta of reconstructed particles in an event, we might see a non-zero transverse momentum. Even though we do not know the exact longitudinal momenta of the initial partons in the event, we do know that the initial transverse momentum is zero. Thus, the residual momenta seen in the transverse plane might be due to particles

which the detector „missed" to reconstruct, which leads us to common definition of missing transverse momentum:

$$\mathbf{p}_T^{\text{miss}} \equiv \not{\mathbf{p}}_T \equiv (\not{p}_x,\, \not{p}_y) = -\sum_{i\in\text{measured}} \mathbf{p}_T^i \,. \tag{B.10}$$

Its magnitude is often referred to as missing transverse energy (MET), $\not{E}_T = |\not{\mathbf{p}}_T|$, or „$p_T$ of MET" because its definition is analogous to (B.7):

$$\not{\mathbf{p}}_T = (\not{E}_T \cos\phi_T^{\text{miss}},\, \not{E}_T \sin\phi_T^{\text{miss}}) \quad \text{with} \quad \phi_T^{\text{miss}} = \arctan\frac{|\not{p}_y|}{|\not{p}_x|} \,. \tag{B.11}$$

The $\not{E}_T$ can be attributed to reconstruction failures, mis-measured momenta or more often neutrinos participating in the event. Another useful quantity akin to MET is MHT defined as the negative vectorial sum of the momenta of high p_T jets (j), leptons (ℓ) and hadronic τ decay products (τ_h):

$$\text{MHT} = |\vec{\not{H}}_T| \quad \text{with} \quad \vec{\not{H}}_T \equiv -\sum_{i\in(j,\ell,\tau_h)} \mathbf{p}_T^i \,. \tag{B.12}$$

The variable $\not{H}_T$ has worse resolution than $\not{E}_T$ but is more robust against PU as it is insensitive to the presence of low p_T hadrons in the event that are characteristic to PU.

B.3. Isolation

The angular separation between two vectors $(p_i^T,\, \eta_i,\, \phi_i)$ and $(p_j^T,\, \eta_j,\, \phi_j)$ is given by

$$\Delta R_{ij} \equiv \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \equiv \sqrt{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2} \,. \tag{B.13}$$

This quantity is used in the „cleaning" of objects: given angular separation constant R and particle A with momentum $(p_T^A,\, \eta_A,\, \phi_A)$, another object B with momentum $(p_T^B,\, \eta_B,\, \phi_B)$ gets „cleaned" w.r.t particle A if $\Delta R_{AB} < R$. Conversely, another object C with momentum $(p_T^C,\, \eta_C,\, \phi_C)$ is retained if $\Delta R_{AC} > R$. The situation is depicted in fig. B.1(b).

For the purpose of this analysis we require the leptons to be isolated (as discussed in section 4.2). The leptons may be part of hadronic jets and therefore reconstructed as such, as opposed to leptons coming from $W/Z/\tau$ decays. One way to impose such isolation condition is to look at a fixed cone size R around a lepton ℓ and check how much transverse momentum is not carried by the lepton defining the cone:

$$I_{\text{iso}}(\ell) = \frac{1}{p_T(\ell)} \sum_{\text{hadrons},\gamma\in R(\ell)} p_T \,, \tag{B.14}$$

where the sum runs over hadrons and photons. If the ratio (B.14) exceeds some predefined value, then we say that the lepton is not isolated.

The above isolation requirement may be sub-optimal in some analyses, however, because it does not account for the situation in which a high p_T top quark decays via $t \to W^+ b$, where the b quark hadronizes into a jet and the W^+ boson decays leptonically. If the top quark has large momentum, the lepton and the b-jet may be highly collimated

in the laboratory frame (i.e. almost parallel to each other other) due to the Lorentz boost in the direction of the top quark.

The remedy is to use a cone of size that varies with the p_T of the lepton:

$$R(\ell) = \frac{10\,\text{GeV}}{\min\left\{\max\left\{p_T(\ell),\ 50\,\text{GeV}\right\},\ 200\,\text{GeV}\right\}}\,. \tag{B.15}$$

The functional relation is plotted in fig. B.2. Efforts in finding the optimal lepton isolation discriminant have resulted in the so-called „mini-isolation", defined by

$$I_{\text{mini}}(\ell) = \frac{1}{p_T(\ell)}\left[\sum_{\text{charged}\in R(\ell)} p_T + \max\left\{0,\ \sum_{\text{neutral}\in\Delta R(\ell)} p_T - \rho\mathcal{A}(\ell)\left(\frac{R(\ell)}{0.3}\right)^2\right\}\right]\,, \tag{B.16}$$

where the first sum extends over charged hadrons, and the second sum over neutral hadrons and photons [49]. The neutral sum incorporates corrections to mitigate PU contributions to the lepton isolation cone. The symbols ρ and $\mathcal{A}(\ell)$ represent the average transverse momentum per unit area and the effective area of the lepton isolation cone. Both ρ and $\mathcal{A}(\ell)$ vary as function of η. Computation of the former is described in [118, 119], whereas the latter is determined from simulations [49]. Its precise values for various η ranges are listed in table D.7.

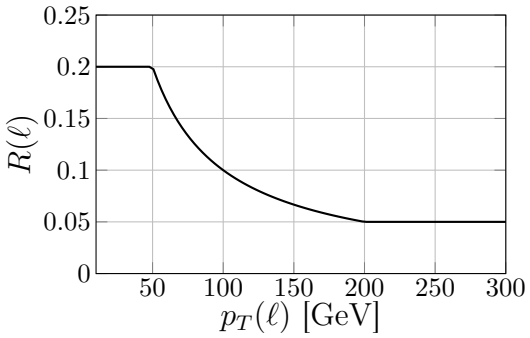


Figure B.2: Size R of the lepton isolation cone as function of lepton p_T.

B.4. Vertexing

The reconstruction of pp collision vertices, which are called „primary" vertices (PVs), is carried out in three steps: selection of the reconstructed tracks, track clustering and vertex fitting to associated tracks [120]. Per bunch crossing multiple (on average 20-25 at current luminosity/PU conditions) PVs are reconstructed, but we are only interested in the PV of the hard scattering process. It is custom to define such PV as the vertex with the largest p_T^2 sum of all its associated tracks because tracks from PU have usually low transverse momentum. The PV is required to be reconstructed within 24 cm in longitudinal direction and 2 cm in transverse direction from the nominal pp interaction point of the detector.

By extrapolating a track towards its PV it is possible to derive various impact parameters (IPs) w.r.t the PV. The transverse IP d_{xy} (or $\text{IP}_{2\text{D}}$ in some texts) is defined as the closest distance of the track to the PV projected onto the transverse plane; the longitudinal IP d_z is the longitudinal distance between the PV and the intercept of the

track with the z-axis; the three-dimensional IP is the closest approach of the track to the PV in three dimensions. These quantities are signed: if the angle α between the vector pointing from the PV to the point of closest approach (also called decay position of the particle origination from the PV) and the track direction is less (greater) than $\pi/2$, then the sign is positive (negative). These definition can be better understood from fig. B.3. The significance of the IP (SIP) is defined as the ratio of the IP to its uncertainty.

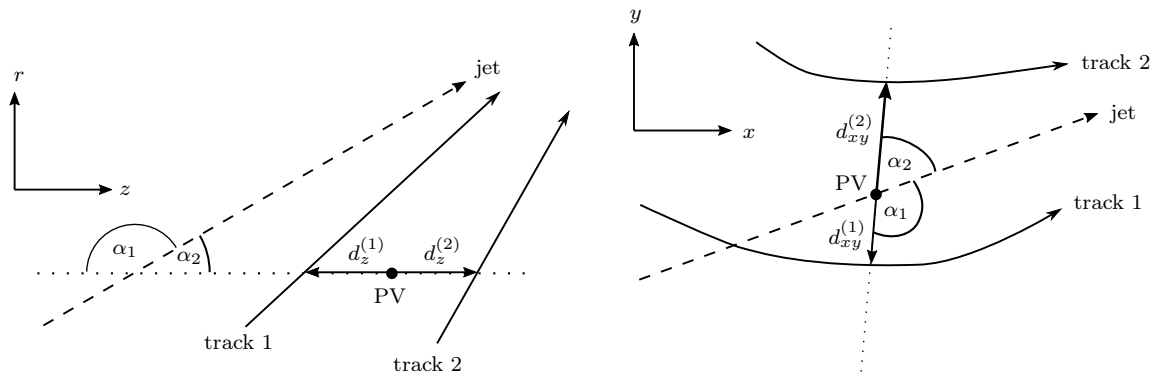


Figure B.3: Impact parameters: longitudinal (left) and transverse (right). Jet direction is indicated with a dashed line, reconstructed tracks shown as solid. Because $\alpha_1 > \pi/2$ ($\alpha_2 < \pi/2$), the corresponding impact parameters $d_z^{(1)}$ and $d_{xy}^{(1)}$ are negative ($d_z^{(2)}$ and $d_{xy}^{(2)}$ are positive).

If a short-lived particle originating from the PV decays into a charged particle, the track of its charged decay product is expected to be close to the PV, meaning that the IPs (and their significances) should be relatively small. This requirement effectively suppresses tracks from PU and from cosmic background. The situation is a bit different for b meson decays. The lifetime of b mesons is sufficiently high for the b mesons to travel a sizable distance before decaying. Its decay position, also called „secondary" vertex (SV), is displaced from the PV by a measurable distance (typically a few mm). A great deal of effort has been invested to reconstruct such SVs and thereby identify b-jets. Out of many b-tagging algorithms, the most popular one is the Combined Secondary Vertex (CSV) algorithm, which exploits characteristic b-jet properties such as distance between the PV and SV; $\mathrm{SIP_{2D}}$ and $\mathrm{SIP_{3D}}$ which tend to take positive values; larger mass associated with the SV; higher track multiplicity; presence of soft leptons within the jet etc [121]. The algorithm feeds this information to a BDT which then computes the CSV b-tagging discriminant, a value in the range of 0 (given jet is not a b-jet) and 1 (given jet is a b-jet).

C. Analysis details

C.1. Fake factor method

This section gives a more explicit derivation of the fake factor (FF) formula (4.8) than what has been published in the original $t\bar{t}h$ multilepton analysis [102, section 7.5]. We denote the efficiency for prompt leptons by ε and the misidentification rate for non-prompt leptons by f. The number of prompt and non-prompt leptons which pass the fakeable (tight) lepton selection criteria are denoted by N_1 ($N_{p|1}$) and N_0 ($N_{p|0}$), respectively. The number of leptons passing the fakeable and tight selection criteria are related to the efficiency and misidentification rate by

$$\varepsilon = \frac{N_{p|1}}{N_1} \quad \text{and} \quad f = \frac{N_{p|0}}{N_0}\,, \tag{C.1}$$

where $N_{p|1}$ ($N_{p|0}$) represents the number of fakeable (non-)prompt leptons that pass the tight selection criteria and N_1 (N_0) the total number of fakeable (non-)prompt leptons. Similarly, we define complementary efficiencies and misidentification rates for the fakeable leptons to fail the tight cuts as

$$\bar{\varepsilon} \overset{!}{=} 1 - \varepsilon = \frac{N_{f|1}}{N_1} \quad \text{and} \quad \bar{f} \overset{!}{=} 1 - f = \frac{N_{f|0}}{N_0}\,, \tag{C.2}$$

where $N_{f|1}$ ($N_{f|0}$) refers to the number of fakeable (non-)prompt leptons that fail the tight selection. This of course implies that

$$N_1 = N_{f|1} + N_{p|1} \quad \text{and} \quad N_0 = N_{f|0} + N_{p|0}\,. \tag{C.3}$$

Now let us consider events which contain a single fakeable lepton. The number of events in which a fakeable lepton that passes (fails) the tight selection reads:

$$N_p = N_{p|1} + N_{p|0} \quad (N_f = N_{f|1} + N_{f|0})\,. \tag{C.4}$$

The efficiencies and fake rates are parametrized by lepton p_T, $|\eta|$ and lepton flavor, but we have dropped these dependencies for the sake of brevity.

We can construct a matrix which relates the total number of events containing prompt and non-prompt leptons (N_1 and N_0) to the number of events in which the fakeable lepton passes or fails the tight lepton selection criteria (N_p and N_f):

$$\begin{pmatrix} N_p \\ N_f \end{pmatrix} = \begin{pmatrix} \varepsilon & f \\ \bar{\varepsilon} & \bar{f} \end{pmatrix} \begin{pmatrix} N_1 \\ N_0 \end{pmatrix}\,. \tag{C.5}$$

We denote the vector on the left hand side by I, the vector on the right hand side by P, and the matrix by M.

We can solve the eq. (C.5) for the vector P by simply inverting the matrix M. However, if we want to assess the contribution of non-prompt leptons $N_{p|0}$ to the signal region (SR), then we have to subtract the contribution of prompt leptons $N_{p|1}$:

$$N_p^{\text{bkg}} \equiv N_p - N_{p|1} \overset{\text{(C.4)}}{=} N_{p|0} \overset{\text{(C.1)}}{=} f N_0\,. \tag{C.6}$$

The number of events with non-prompt leptons that pass the tight lepton selection criteria represents the fake lepton background in the SR, which we denote by N_p^{bkg}. By solving the eq. (C.5) for N_0, the contribution of non-prompt leptons to the SR reads

$$N_p^{\text{bkg}} = \frac{f}{\varepsilon\bar{f} - \bar{\varepsilon}f}(\varepsilon N_f - \bar{\varepsilon}N_p)\,. \tag{C.7}$$

In practice the efficiency for prompt leptons to pass the tight cuts is greater than that of non-prompt leptons ($\varepsilon > f$), which implies that $\bar{f} > \bar{\varepsilon}$ and $\varepsilon\bar{f} > f\bar{\varepsilon}$. Furthermore, the background usually contains fewer events with prompt leptons than events with non-prompt leptons ($N_0 > N_1$). Therefore, $\bar{\varepsilon}N_1$ is small compared to $\bar{f}N_0$ and we can make an approximation $N_f \approx fN_0$ from which

$$N_p^{\text{bkg}} \approx \frac{f}{\bar{f}}N_f = FN_f\,, \tag{C.8}$$

where we have defined $F = f/\bar{f}$. The same result can be obtained by setting $\bar{\varepsilon} = 0$ in (C.5):

$$\widetilde{M} \equiv M|_{\bar{\varepsilon}=0} = \begin{pmatrix} \varepsilon & f \\ 0 & \bar{f} \end{pmatrix}\,. \tag{C.9}$$

This matrix serves as a building block in the FF method that is used to estimate the fake background, as we will see shortly.

Let us now consider the case of events that contain two fakeable leptons. If we try to infer the fakeable background in the SR in this case, we can categorize the sample of events into four categories that are based on the multiplicity of prompt and non-prompt leptons: both leptons are prompt (N_{11}); one lepton is prompt, the other is non-prompt (N_{10}, N_{01}); or both leptons are non-prompt (N_{00}). An alternative is to categorize the sample of events by the number of leptons that pass the tight lepton selection: both pass (N_{pp}), one fails and the other does not (N_{pf}, N_{fp}), or both fail (N_{ff}). For the sake of consistence, we assume that the i-th index in both categorization options refers to i-th lepton sorted by their p_T in decreasing order, e.g. N_{fp} is the number of events in which the leading fakeable lepton fails but the subleading one passes the tight cuts. An equation equivalent to (C.5) is obtained by taking the tensor product of (C.5):

$$I_1 \otimes I_2 = (M_1 \otimes M_2)(P_1 \otimes P_2)\,. \tag{C.10}$$

With

$$M_1 \otimes M_2 = \begin{pmatrix} \varepsilon_1 \begin{pmatrix} \varepsilon_2 & f_2 \\ \bar{\varepsilon}_2 & \bar{f}_2 \end{pmatrix} & f_1 \begin{pmatrix} \varepsilon_2 & f_2 \\ \bar{\varepsilon}_2 & \bar{f}_2 \end{pmatrix} \\ \bar{\varepsilon}_1 \begin{pmatrix} \varepsilon_2 & f_2 \\ \bar{\varepsilon}_2 & \bar{f}_2 \end{pmatrix} & \bar{f}_1 \begin{pmatrix} \varepsilon_2 & f_2 \\ \bar{\varepsilon}_2 & \bar{f}_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \varepsilon_1\varepsilon_2 & \varepsilon_1 f_2 & f_1\varepsilon_2 & f_1 f_2 \\ \varepsilon_1\bar{\varepsilon}_2 & \varepsilon_1\bar{f}_2 & f_1\bar{\varepsilon}_2 & f_1\bar{f}_2 \\ \bar{\varepsilon}_1\varepsilon_2 & \bar{\varepsilon}_1 f_2 & \bar{f}_1\varepsilon_2 & \bar{f}_1 f_2 \\ \bar{\varepsilon}_1\bar{\varepsilon}_2 & \bar{\varepsilon}_1\bar{f}_2 & \bar{f}_1\bar{\varepsilon}_2 & \bar{f}_1\bar{f}_2 \end{pmatrix}\,, \tag{C.11}$$

the equation (C.10) takes the following explicit form:

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \begin{pmatrix} \varepsilon_1\varepsilon_2 & \varepsilon_1 f_2 & f_1\varepsilon_2 & f_1 f_2 \\ \varepsilon_1\bar{\varepsilon}_2 & \varepsilon_1\bar{f}_2 & f_1\bar{\varepsilon}_2 & f_1\bar{f}_2 \\ \bar{\varepsilon}_1\varepsilon_2 & \bar{\varepsilon}_1 f_2 & \bar{f}_1\varepsilon_2 & \bar{f}_1 f_2 \\ \bar{\varepsilon}_1\bar{\varepsilon}_2 & \bar{\varepsilon}_1\bar{f}_2 & \bar{f}_1\bar{\varepsilon}_2 & \bar{f}_1\bar{f}_2 \end{pmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix}\,. \tag{C.12}$$

The contribution of non-prompt leptons to the SR can of course be estimated by inverting the above matrix and solving for the vector $P_1 \otimes P_2$, however, the exact solution already becomes quite elaborate. Instead, we use the approximation $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = 0$ or, in analogy to (C.9), the tensor product of the approximate matrices $(\widetilde{M}_1 \otimes \widetilde{M}_2)$:

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \begin{pmatrix} \varepsilon_1\varepsilon_2 & \varepsilon_1 f_2 & f_1\varepsilon_2 & f_1 f_2 \\ 0 & \varepsilon_1 \bar{f}_2 & 0 & f_1 \bar{f}_2 \\ 0 & 0 & \bar{f}_1\varepsilon_2 & \bar{f}_1 f_2 \\ 0 & 0 & 0 & \bar{f}_1 \bar{f}_2 \end{pmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix} . \tag{C.13}$$

We can now solve for the fake background yield N_{pp}^{bkg} by the means of expressions:

$$N_{ff} = \bar{f}_1 f_2 N_{00} \;\Rightarrow\; N_{00} = \frac{N_{ff}}{\bar{f}_1 \bar{f}_2} , \tag{C.14}$$

$$N_{fp} = \bar{f}_1 \varepsilon_2 N_{01} + \bar{f}_1 f_2 N_{00} \;\Rightarrow\; N_{01} = \frac{1}{\bar{f}_1 \varepsilon_2}(N_{fp} - F_2 N_{ff}) , \tag{C.15}$$

$$N_{pf} = \varepsilon_1 \bar{f}_2 + f_1 \bar{f}_2 N_{00} \;\Rightarrow\; N_{10} = \frac{1}{\varepsilon_1 \bar{f}_2}(N_{pf} - F_1 N_{ff}) , \tag{C.16}$$

$$\begin{aligned} N_{pp}^{\text{bkg}} &= \varepsilon_1 f_2 N_{10} + f_1 \varepsilon_2 N_{01} + f_1 f_2 N_{00} = \\ &= F_2(N_{pf} - F_1 N_{ff}) + F_1(N_{fp} - F_2 N_{ff}) + F_1 F_2 N_{ff} = \\ &= F_1 N_{fp} + F_2 N_{pf} - F_1 F_2 N_{ff} . \end{aligned} \tag{C.17}$$

Here the fake factor F_i applies to events in which the i-th fakeable lepton fails the tight selection.

The procedure for the case of events that contain three fakeable leptons is analogous: taking the triple tensor product of (C.5),

$$I_1 \otimes I_2 \otimes I_3 = (M_1 \otimes M_2 \otimes M_3)(P_1 \otimes P_2 \otimes P_3) , \tag{C.18}$$

we obtain

$$\begin{pmatrix} N_{ppp} \\ N_{ppf} \\ N_{pfp} \\ N_{pff} \\ N_{fpp} \\ N_{fpf} \\ N_{ffp} \\ N_{fff} \end{pmatrix} = \begin{pmatrix} \varepsilon_1\varepsilon_2\varepsilon_3 & \varepsilon_1\varepsilon_2 f_3 & \varepsilon_1 f_2\varepsilon_3 & \varepsilon_1 f_2 f_3 & f_1\varepsilon_2\varepsilon_3 & f_1\varepsilon_2 f_3 & f_1 f_2\varepsilon_3 & f_1 f_2 f_3 \\ \varepsilon_1\varepsilon_2\bar{\varepsilon}_3 & \varepsilon_1\varepsilon_2 \bar{f}_3 & \varepsilon_1 f_2\bar{\varepsilon}_3 & \varepsilon_1 f_2 \bar{f}_3 & f_1\varepsilon_2\bar{\varepsilon}_3 & f_1\varepsilon_2 \bar{f}_3 & f_1 f_2\bar{\varepsilon}_3 & f_1 f_2 \bar{f}_3 \\ \varepsilon_1\bar{\varepsilon}_2\varepsilon_3 & \varepsilon_1\bar{\varepsilon}_2 f_3 & \varepsilon_1 \bar{f}_2\varepsilon_3 & \varepsilon_1 \bar{f}_2 f_3 & f_1\bar{\varepsilon}_2\varepsilon_3 & f_1\bar{\varepsilon}_2 f_3 & f_1 \bar{f}_2\varepsilon_3 & f_1 \bar{f}_2 f_3 \\ \varepsilon_1\bar{\varepsilon}_2\bar{\varepsilon}_3 & \varepsilon_1\bar{\varepsilon}_2 \bar{f}_3 & \varepsilon_1 \bar{f}_2\bar{\varepsilon}_3 & \varepsilon_1 \bar{f}_2 \bar{f}_3 & f_1\bar{\varepsilon}_2\bar{\varepsilon}_3 & f_1\bar{\varepsilon}_2 \bar{f}_3 & f_1 \bar{f}_2\bar{\varepsilon}_3 & f_1 \bar{f}_2 \bar{f}_3 \\ \bar{\varepsilon}_1\varepsilon_2\varepsilon_3 & \bar{\varepsilon}_1\varepsilon_2 f_3 & \bar{\varepsilon}_1 f_2\varepsilon_3 & \bar{\varepsilon}_1 f_2 f_3 & \bar{f}_1\varepsilon_2\varepsilon_3 & \bar{f}_1\varepsilon_2 f_3 & \bar{f}_1 f_2\varepsilon_3 & \bar{f}_1 f_2 f_3 \\ \bar{\varepsilon}_1\varepsilon_2\bar{\varepsilon}_3 & \bar{\varepsilon}_1\varepsilon_2 \bar{f}_3 & \bar{\varepsilon}_1 f_2\bar{\varepsilon}_3 & \bar{\varepsilon}_1 f_2 \bar{f}_3 & \bar{f}_1\varepsilon_2\bar{\varepsilon}_3 & \bar{f}_1\varepsilon_2 \bar{f}_3 & \bar{f}_1 f_2\bar{\varepsilon}_3 & \bar{f}_1 f_2 \bar{f}_3 \\ \bar{\varepsilon}_1\bar{\varepsilon}_2\varepsilon_3 & \bar{\varepsilon}_1\bar{\varepsilon}_2 f_3 & \bar{\varepsilon}_1 \bar{f}_2\varepsilon_3 & \bar{\varepsilon}_1 \bar{f}_2 f_3 & \bar{f}_1\bar{\varepsilon}_2\varepsilon_3 & \bar{f}_1\bar{\varepsilon}_2 f_3 & \bar{f}_1 \bar{f}_2\varepsilon_3 & \bar{f}_1 \bar{f}_2 f_3 \\ \bar{\varepsilon}_1\bar{\varepsilon}_2\bar{\varepsilon}_3 & \bar{\varepsilon}_1\bar{\varepsilon}_2 \bar{f}_3 & \bar{\varepsilon}_1 \bar{f}_2\bar{\varepsilon}_3 & \bar{\varepsilon}_1 \bar{f}_2 \bar{f}_3 & \bar{f}_1\bar{\varepsilon}_2\bar{\varepsilon}_3 & \bar{f}_1\bar{\varepsilon}_2 \bar{f}_3 & \bar{f}_1 \bar{f}_2\bar{\varepsilon}_3 & \bar{f}_1 \bar{f}_2 \bar{f}_3 \end{pmatrix} \begin{pmatrix} N_{111} \\ N_{110} \\ N_{101} \\ N_{100} \\ N_{011} \\ N_{010} \\ N_{001} \\ N_{000} \end{pmatrix} . \tag{C.19}$$

The above expression simplifies once the assumptions $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = \bar{\varepsilon}_3 = 0$ are made or, in analogy to (C.9), the tensor product of the approximate matrices $(\widetilde{M}_1 \otimes \widetilde{M}_2 \otimes \widetilde{M}_3)$ is used:

$$\begin{pmatrix} N_{ppp} \\ N_{ppf} \\ N_{pfp} \\ N_{pff} \\ N_{fpp} \\ N_{fpf} \\ N_{ffp} \\ N_{fff} \end{pmatrix} = \begin{pmatrix} \varepsilon_1\varepsilon_2\varepsilon_3 & \varepsilon_1\varepsilon_2 f_3 & \varepsilon_1 f_2\varepsilon_3 & \varepsilon_1 f_2 f_3 & f_1\varepsilon_2\varepsilon_3 & f_1\varepsilon_2 f_3 & f_1 f_2\varepsilon_3 & f_1 f_2 f_3 \\ 0 & \varepsilon_1\varepsilon_2 \bar{f}_3 & 0 & \varepsilon_1 f_2 \bar{f}_3 & 0 & f_1\varepsilon_2 \bar{f}_3 & 0 & f_1 f_2 \bar{f}_3 \\ 0 & 0 & \varepsilon_1 \bar{f}_2\varepsilon_3 & \varepsilon_1 \bar{f}_2 f_3 & 0 & 0 & f_1 \bar{f}_2\varepsilon_3 & f_1 \bar{f}_2 f_3 \\ 0 & 0 & 0 & \varepsilon_1 \bar{f}_2 \bar{f}_3 & 0 & 0 & 0 & f_1 \bar{f}_2 \bar{f}_3 \\ 0 & 0 & 0 & 0 & \bar{f}_1\varepsilon_2\varepsilon_3 & \bar{f}_1\varepsilon_2 f_3 & \bar{f}_1 f_2\varepsilon_3 & \bar{f}_1 f_2 f_3 \\ 0 & 0 & 0 & 0 & 0 & \bar{f}_1\varepsilon_2 \bar{f}_3 & 0 & \bar{f}_1 f_2 \bar{f}_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bar{f}_1 \bar{f}_2\varepsilon_3 & \bar{f}_1 \bar{f}_2 f_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{f}_1 \bar{f}_2 \bar{f}_3 \end{pmatrix} \begin{pmatrix} N_{111} \\ N_{110} \\ N_{101} \\ N_{100} \\ N_{011} \\ N_{010} \\ N_{001} \\ N_{000} \end{pmatrix} . \tag{C.20}$$

This enables us to solve for the number of fake background events in the SR:

$$N_{fff} = \bar{f}_1\bar{f}_2\bar{f}_3 N_{000} \;\Rightarrow\; N_{000} = \frac{N_{fff}}{\bar{f}_1\bar{f}_2\bar{f}_3}\,, \tag{C.21}$$

$$N_{pff} = \varepsilon_1\bar{f}_2\bar{f}_3 N_{100} + f_1\bar{f}_2\bar{f}_3 N_{000} \;\Rightarrow\; N_{100} = \frac{N_{pff} - F_1 N_{fff}}{\varepsilon_1\bar{f}_2\bar{f}_3}\,, \tag{C.22}$$

$$N_{fpf} = \bar{f}_1\varepsilon_2\bar{f}_3 N_{010} + \bar{f}_1 f_2\bar{f}_3 N_{000} \;\Rightarrow\; N_{010} = \frac{N_{fpf} - F_2 N_{fff}}{\bar{f}_1\varepsilon_2\bar{f}_3}\,, \tag{C.23}$$

$$N_{ffp} = \bar{f}_1\bar{f}_2\varepsilon_3 N_{001} + \bar{f}_1\bar{f}_2 f_3 N_{000} \;\Rightarrow\; N_{001} = \frac{N_{ffp} - F_3 N_{fff}}{\bar{f}_1\bar{f}_2\varepsilon_3}\,, \tag{C.24}$$

$$\begin{aligned} N_{ppf} &= \varepsilon_1\varepsilon_2\bar{f}_3 N_{110} + \varepsilon_1 f_2\bar{f}_3 N_{100} + f_1\varepsilon_2\bar{f}_3 N_{010} + f_1 f_2\bar{f}_3 N_{000} = \\ &= \varepsilon_1\varepsilon_2\bar{f}_3 N_{110} + F_2(N_{pff} - F_1 N_{fff}) + F_1(N_{fpf} - F_2 N_{fff}) + F_1F_2 N_{fff} \;\Rightarrow \\ \Rightarrow\; N_{110} &= \frac{1}{\varepsilon_1\varepsilon_2\bar{f}_3}(N_{ppf} - F_2 N_{pff} - F_1 N_{fpf} + F_1F_2 N_{fff})\,, \end{aligned} \tag{C.25}$$

$$\begin{aligned} N_{pfp} &= \varepsilon_1\bar{f}_2\varepsilon_3 N_{101} + \varepsilon_1\bar{f}_2 f_3 N_{100} + f_1\bar{f}_2\varepsilon_3 N_{001} + f_1\bar{f}_2 f_3 N_{000} = \\ &= \varepsilon_1\bar{f}_2\varepsilon_3 N_{101} + F_3(N_{pff} - F_1 N_{fff}) + F_1(N_{ffp} - F_3 N_{fff}) + F_1F_3 N_{fff} \;\Rightarrow \\ \Rightarrow\; N_{101} &= \frac{1}{\varepsilon_1\bar{f}_2\varepsilon_3}(N_{pfp} - F_3 N_{pff} - F_1 N_{ffp} + F_1F_3 N_{fff})\,, \end{aligned} \tag{C.26}$$

$$\begin{aligned} N_{fpp} &= \bar{f}_1\varepsilon_2\varepsilon_3 N_{011} + \bar{f}_1\varepsilon_2 f_3 N_{010} + \bar{f}_1 f_2\varepsilon_3 N_{001} + \bar{f}_1 f_2 f_3 N_{000} = \\ &= \bar{f}_1\varepsilon_2\varepsilon_3 N_{011} + F_3(N_{fpf} - F_2 N_{fff}) + F_2(N_{ffp} - F_3 N_{fff}) + F_2F_3 N_{fff} \;\Rightarrow \\ \Rightarrow\; N_{011} &= \frac{1}{\bar{f}_1\varepsilon_2\varepsilon_3}(N_{fpp} - F_3 N_{fpf} - F_2 N_{ffp} + F_2F_3 N_{fff})\,, \end{aligned} \tag{C.27}$$

$$\begin{aligned} N^{\text{bkg}}_{ppp} &= \varepsilon_1\varepsilon_2 f_3 N_{110} + \varepsilon_1 f_2\varepsilon_3 N_{101} + f_1\varepsilon_2\varepsilon_3 N_{011} + \\ &\quad + f_1 f_2\varepsilon_3 N_{001} + f_1\varepsilon_2 f_3 N_{010} + \varepsilon_1 f_2 f_3 N_{100} + f_1 f_2 f_3 N_{000} = \\ &= F_3(N_{ppf} - F_2 N_{pff} - F_1 N_{fpf} + F_1F_2 N_{fff}) + \\ &\quad + F_2(N_{pfp} - F_3 N_{pff} - F_1 N_{ffp} + F_1F_3 N_{fff}) + \\ &\quad + F_1(N_{fpp} - F_3 N_{fpf} - F_2 N_{ffp} + F_2F_3 N_{fff}) + \\ &\quad + F_2F_3(N_{pff} - F_1 N_{fff}) + F_1F_3(N_{fpf} - F_2 N_{fff}) + F_1F_2(N_{ffp} - F_3 N_{fff}) + \\ &\quad + F_1F_2F_3 N_{fff} = \end{aligned} \tag{C.28}$$

$$\begin{aligned} &= F_1 N_{fpp} + F_2 N_{pfp} + F_3 N_{ppf} - F_1F_2 N_{ffp} - F_1F_3 N_{fpf} - F_2F_3 N_{pff} + \\ &\quad + F_1F_2F_3 N_{fff}\,. \end{aligned} \tag{C.29}$$

The generalization to the case of events containing n leptons is evident: just perform n tensor products of (C.5), while assuming (C.9), and solve for $N^{\text{bkg}}_{p^n}$. The general FF solution exhibits a recursive structure and may be symbolically written as

$$N^{\text{bkg}}_{p^n} = \sum_{k=1}^{n}(-1)^{k+1} \prod_{j\in\{i_1,\ldots,i_k\}} F_{i_j} N_{\text{perm}_j(f^k p^{n-k})}\,, \tag{C.30}$$

where the product extends over all possible combinations of k leptons that fail the tight lepton selection out of n leptons in total.

It is important to stress that each event in the application region is assigned a unique FF computed as function of lepton p_T, η and lepton flavor. The total fake yield N^{bkg}

is obtained by summing all such weights together. The expressions in (C.8), (C.17) and (C.29) then generalize to

$$N_{p}^{\text{bkg}} = \sum_{f} F_1 \,, \tag{C.31}$$

$$N_{pp}^{\text{bkg}} = \sum_{fp} F_1 + \sum_{pf} F_2 - \sum_{ff} F_1 F_2 \,, \tag{C.32}$$

$$N_{ppp}^{\text{bkg}} = \sum_{fpp} F_1 + \sum_{pfp} F_2 + \sum_{ppf} F_3 - \sum_{ffp} F_1 F_2 - \sum_{fpf} F_1 F_3 - \sum_{pff} F_2 F_3 + \sum_{fff} F_1 F_2 F_3 \,, \tag{C.33}$$

where the sums extend over events containing either one, two or three fakeable leptons. The fake factors F_i are computed from the p_T, η and lepton flavor of the i-th leading lepton that passes the fakeable lepton selection criteria, but fails the tight lepton selection criteria. Only these leptons contribute to the sums (C.31)-(C.33).

C.2. Statistical treatment

The results presented in section 4.4.4 are derived by the means of a maximum likelihood (ML) fit. The ATLAS and CMS collaborations have adapted and achieved convergence on the method so that there would be no ambiguity in presenting the results [122]. In practice, the statistical procedures are carried out using the RooStats-based [123] statistical analysis software packages CombinedLimit [124] and CombineHarvester [125].

Essentially, the method promotes the signal strength μ (given by eq. 2.38) to be the parameter of interest (POI) for the purpose of performing hypotheses tests. The null hypothesis H_0 is that our measured data contains only background, which corresponds to $\mu = 0$. Since we actually do not know what part of the measured data is signal and which part is background, the alternative hypothesis H_1 is formulated as signal-plus-background hypothesis, corresponding to $\mu = 1$.

Given a hypothesis H_μ and binned distributions of the final discriminant for data, the $t\bar{t}h$ signal and different backgrounds, we would expect to see

$$p_i(\theta) = b_i(\theta) + \mu s_i(\theta) \tag{C.34}$$

events in the i-th bin of the discriminant distribution, where s_i (b_i) denotes the number of signal (background) events in the bin. The set of nuisance parameters θ represent systematic uncertainties that may scale the event yield, change the shape of the discriminant distribution, or both. Given eq. (C.34), the probability to observe d_i data events in the i-th bin of the discriminant distribution is given by the Poisson distribution:

$$\mathcal{L}_i(\mu,\, \theta) = e^{-p_i} \frac{p_i^{d_i}}{d_i!} \,. \tag{C.35}$$

The number of events observed in each bin is statistically independent from the number of events observed in other bins. The overall probability to observe a given distribution of the discriminant, specified by the N values of d_i, is thus given by the product of (C.35) over all N bins:

$$\mathcal{L}(\mu,\, \theta) = \prod_{i=1}^{N} \mathcal{L}_i \times p(\widetilde{\theta}|\theta) = \prod_{i=1}^{N} e^{-b_i(\theta) - \mu s_i(\theta)} \frac{[b_i(\theta) + \mu s_i(\theta)]^{d_i}}{d_i!} \times p(\widetilde{\theta}|\theta) \,. \tag{C.36}$$

In (C.36), we have also included a probability distribution function (p.d.f) for the systematic uncertainties. The function $p(\widetilde{\theta}|\theta)$ reflects our degree of belief on what the true value of a nuisance parameter θ is, given a measured value $\widetilde{\theta}$. The exact form of $p(\widetilde{\theta}|\theta)$ depends on the nature of systematic uncertainties, e.g. shape-changing systematics follow Gaussian p.d.f, statistical uncertainties are described with Gamma distribution, and other yield-changing systematics are modeled with log-normal distribution [122].

The likelihood function $\mathcal{L}(\hat{\mu},\, \hat{\theta})$ is called unconditional maximum likelihood, if its arguments

$$(\hat{\mu},\, \hat{\theta}) = \arg\max_{\mu,\,\theta} \mathcal{L}(\mu,\, \theta) \tag{C.37}$$

maximize the likelihood function (C.36). The value $\hat{\mu}$ represents the best possible fit of the signal strength, oftentimes referred to as the observed signal strength. By letting the signal strength float freely in (C.37),

$$\bar{\theta}_\mu = \arg\max_{\theta} \mathcal{L}(\mu,\, \theta)\,, \tag{C.38}$$

we obtain the maximum conditional likelihood $\mathcal{L}(\mu,\, \bar{\theta}_\mu)$. Notice that the nuisance parameters $\bar{\theta}_\mu$ which maximize the likelihood function (C.36) depend on the given value of μ which is indicated by subscript.

If we were to plot $-2\ln\mathcal{L}(\mu,\, \bar{\theta}_\mu)$ as a function of μ, the best fit value $\hat{\mu}$ would correspond to a minimum. The plotted curve is parabola-like around the minimum, and becomes broader as the systematic uncertainties increase, reflecting a loss of information caused by uncertainties. The systematic and statistical uncertainties $\mu_\pm$ that are associated with the best fit value $\hat{\mu}$ are obtained by moving one unit up from the minimum and finding the abscissa of the horizontal intersection to the fit. The best fit is often quoted as $\hat{\mu}^{\delta_+}_{\delta_-}$, where the uncertainties are expressed via differences $\delta_\pm = \mu_\pm - \hat{\mu}$.

During the statistical procedures we have avoided to look at the data – often referred to as „blind analysis" – in order to resist the temptation of fine-tuning the analysis and therefore produce subjective results. Furthermore, focusing only on the observation does not give us any metric for comparison. This has motivated to derive expected statistical figures analogous to the observed ones. The expected values of the statistical figures are obtained by replacing the real data with „pseudo-data", also known as Asimov data, which are constructed by randomly sampling from the expected numbers p_i given by (C.34), assuming either the background-only ($\mu = 0$) or the signal-plus-background ($\mu = 1$) hypothesis. The nuisance parameters θ are randomly sampled from the p.d.f $p(\tilde{\theta}|\theta)$. All relevant statistical quantities are recalculated from this pseudo-data, and quoted as „expected" instead of „observed" or „measured" results. The expected values serve as a guideline to quantify the sensitivity of an analysis.

It is useful to construct a profile likelihood function as the ratio of unconditional over the conditional likelihood,

$$\lambda(\mu) = \frac{\mathcal{L}(\mu,\, \bar{\theta})}{\mathcal{L}(\hat{\mu},\, \hat{\theta})}\,. \tag{C.39}$$

The likelihood given by (C.36) is by construction positive semi-definite, which in turn implies that $0 \leqslant \lambda(\mu) \leqslant 1$. If the profile likelihood $\lambda(\mu)$ takes values close to 1 for a given

hypothesis H_μ, then the hypothesis agrees well with the observation (i.e. the data). In practice it is useful to express the level of agreement between hypothesis and data by a modified test statistic

$$q_\mu = -2 \ln \lambda(\mu) \,, \tag{C.40}$$

which has the advantage of turning the factors in the expression for the likelihood in (C.36) into sums. With this definition, smaller (larger) values of the test statistic indicate a higher (lower) level of agreement between a given hypothesis and the data.

The test statistic (C.40) helps to assess how likely it is for an excess that may be observed in the data to arise from a statistical fluctuations of the background, and allows to compute the probability, also known as significance, for this to happen. In order to obtain the significance, let us consider the following test statistic for rejecting the background-only hypothesis ($\mu = 0$):

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \text{if } \hat{\mu} \geqslant 0 \,, \\ 0 & \text{if } \hat{\mu} < 0 \,. \end{cases} \tag{C.41}$$

Any positive signal ($\hat{\mu} > 0$) implies incompatibility with the null hypothesis, whereas a deficit in the data w.r.t the background-only hypothesis indicates that either the background is poorly modeled or the observed data is due to a downward fluctuation, and therefore not interpreted as evidence for the presence of a signal. Assuming that the null hypothesis and the corresponding nuisance parameters $\bar{\theta}_0^{\text{obs}}$ are obtained from a fit to the data, the distribution of q_0 follows the p.d.f $f(q_0|H_0, \bar{\theta}_0^{\text{obs}})$, which is typically constructed by random sampling from the likelihood function of the background-only hypothesis $\mathcal{L}(0, \bar{\theta}_0^{\text{obs}})$, given by (C.36) with $\mu = 0$. The probability to observe a value of the test statistic as large as q_0^{obs} (or greater), assuming the null hypothesis, is given by p-value

$$p_0 = \int_{q_0^{\text{obs}}}^{\infty} f(q_0|H_0, \bar{\theta}_0^{\text{obs}}) \, \mathrm{d}q_0 \,, \tag{C.42}$$

where q_0^{obs} is the observed value of q_0 which is obtained from the fit to the data. The p-value corresponds to the cumulative probability $f(q_0|H_0, \bar{\theta}_0^{\text{obs}})$ integrated over the region $q_0 \geqslant q_0^{\text{obs}}$. It is convenient to relate the observed value of p_0 to the tail probability of a Gaussian distribution,

$$p_0 = \int_Z^{\infty} \frac{\exp(-x^2/2)}{\sqrt{2\pi}} \, \mathrm{d}x \,, \tag{C.43}$$

where Z denotes the number of standard deviations corresponding to the observed probability p_0 and is commonly called the significance. Intuitively, the larger the significance, the smaller the p-value, or equivalently the probability for an excess observed in the data to arise from a statistical fluctuation of the background.

D. Tables

GGF [pb]	VBF [fb]	Higgs *Strahlung*		$b\bar{b}h$ [fb]	$t\bar{t}h$ [fb]	$th + \bar{t}h$ [fb] (t-channel)	total [pb]
		Wh [fb]	Zh [fb]				
48.57	3925	1373.00	883.70	488.0	507.1	74.25	55.71

Table D.1: The total SM cross sections for the main Higgs production mechanisms at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ for Higgs boson of mass $m_h = 125\,\text{GeV}$ [32, 126]. The cross section for s-channel $th + \bar{t}h$ production is a few orders of magnitude smaller than the cross section for t-channel $th + \bar{t}h$ production and thus not shown here.

	$e^+\nu_e(\bar{\nu}_\tau)^\dagger$	$\mu^+\nu_\mu(\bar{\nu}_\tau)^\dagger$	$\tau^+\nu_\tau$	$\ell^+\nu_\ell(\bar{\nu}_\tau)^\dagger$		hadronic	
				avg	sum	$q\bar{q}'$	$\tau_h^+\bar{\nu}_\tau$
W^+	10.71%	10.63%	11.38%	10.87%	21.34%	67.41%	-
τ^+	17.82%	17.39%	-	35.21%	17.61%	-	64.79%

	e^+e^-	$\mu^+\mu^-$	$\tau^+\tau^-$	$\ell^+\ell^-$		hadronic ($q\bar{q}$)
				avg	sum	
Z	3.363%	3.366%	3.370%	3.365%	6.729%	69.91%

† – the tau neutrino arises only from leptonic τ decay.

Table D.2: Branching ratios (Br) of W, τ and Z [24]. The symbol ℓ stands for either e or μ. $\text{Br}(\tau \to \tau_h\nu_\tau)$ represents all hadronic τ decay modes. See also table D.12.

$h \to$	$\gamma\gamma$	ZZ	W^+W^-	$\tau^+\tau^-$	$b\bar{b}$	$Z\gamma$	$\mu^+\mu^-$
branching ratio $\times 10^{-3}$	$2.27^{+5.0\%}_{-4.9\%}$	$26.2^{+4.3\%}_{-4.1\%}$	$214^{+4.3\%}_{-4.2\%}$	$62.7^{+5.7\%}_{-5.7\%}$	$584^{+3.2\%}_{-3.3\%}$	$1.53^{+9.0\%}_{-8.9\%}$	$0.218^{+6.0\%}_{-5.9\%}$
mass resolution	1-2%	1-2%	20%	15%	10%	-	-

Table D.3: Branching ratios of Higgs decay modes with relative uncertainties and typical resolution on the reconstructed Higgs mass in different decay modes [24, tables 11.3-11.4].

Dataset name	Run-range	Integrated luminosity [fb^{-1}]
/SingleElectron/Run2016B-23Sep2016-v3/MINIAOD	273150–275376	5.79
/SingleElectron/Run2016C-23Sep2016-v1/MINIAOD	275656–276283	2.57
/SingleElectron/Run2016D-23Sep2016-v1/MINIAOD	276315–276811	4.25
/SingleElectron/Run2016E-23Sep2016-v1/MINIAOD	276831–277420	4.01
/SingleElectron/Run2016F-23Sep2016-v1/MINIAOD	277932–278808	3.10
/SingleElectron/Run2016G-23Sep2016-v1/MINIAOD	278820–280385	7.54
/SingleElectron/Run2016H-PromptReco-v2/MINIAOD	281207–284035	8.39
/SingleElectron/Run2016H-PromptReco-v3/MINIAOD	284036–284044	0.22
/SingleMuon/Run2016B-23Sep2016-v3/MINIAOD	273150–275376	5.79
/SingleMuon/Run2016C-23Sep2016-v1/MINIAOD	275656–276283	2.57
/SingleMuon/Run2016D-23Sep2016-v1/MINIAOD	276315–276811	4.25
/SingleMuon/Run2016E-23Sep2016-v1/MINIAOD	276831–277420	4.01
/SingleMuon/Run2016F-23Sep2016-v1/MINIAOD	277932–278808	3.10
/SingleMuon/Run2016G-23Sep2016-v1/MINIAOD	278820–280385	7.54
/SingleMuon/Run2016H-PromptReco-v2/MINIAOD	281207–284035	8.39
/SingleMuon/Run2016H-PromptReco-v3/MINIAOD	284036–284044	0.22
/DoubleEG/Run2016B-23Sep2016-v3/MINIAOD	273150–275376	5.79
/DoubleEG/Run2016C-23Sep2016-v1/MINIAOD	275656–276283	2.57
/DoubleEG/Run2016D-23Sep2016-v1/MINIAOD	276315–276811	4.25
/DoubleEG/Run2016E-23Sep2016-v1/MINIAOD	276831–277420	4.01
/DoubleEG/Run2016F-23Sep2016-v1/MINIAOD	277932–278808	3.10
/DoubleEG/Run2016G-23Sep2016-v1/MINIAOD	278820–280385	7.54
/DoubleEG/Run2016H-PromptReco-v2/MINIAOD	281207–284035	8.39
/DoubleEG/Run2016H-PromptReco-v3/MINIAOD	284036–284044	0.22
/DoubleMuon/Run2016B-23Sep2016-v3/MINIAOD	273150–275376	5.79
/DoubleMuon/Run2016C-23Sep2016-v1/MINIAOD	275656–276283	2.57
/DoubleMuon/Run2016D-23Sep2016-v1/MINIAOD	276315–276811	4.25
/DoubleMuon/Run2016E-23Sep2016-v1/MINIAOD	276831–277420	4.01
/DoubleMuon/Run2016F-23Sep2016-v1/MINIAOD	277932–278808	3.10
/DoubleMuon/Run2016G-23Sep2016-v1/MINIAOD	278820–280385	7.54
/DoubleMuon/Run2016H-PromptReco-v2/MINIAOD	281207–284035	8.39
/DoubleMuon/Run2016H-PromptReco-v3/MINIAOD	284036–284044	0.22
/MuonEG/Run2016B-23Sep2016-v3/MINIAOD	273150–275376	5.79
/MuonEG/Run2016C-23Sep2016-v1/MINIAOD	275656–276283	2.57
/MuonEG/Run2016D-23Sep2016-v1/MINIAOD	276315–276811	4.25
/MuonEG/Run2016E-23Sep2016-v1/MINIAOD	276831–277420	4.01
/MuonEG/Run2016F-23Sep2016-v1/MINIAOD	277932–278808	3.10
/MuonEG/Run2016G-23Sep2016-v1/MINIAOD	278820–280385	7.54
/MuonEG/Run2016H-PromptReco-v2/MINIAOD	281207–284035	8.39
/MuonEG/Run2016H-PromptReco-v3/MINIAOD	284036–284044	0.22
Total	-	35.87

Table D.4: Datasets collected with different triggers that are used in this analysis. The data is masked with `Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt` released by the CMS DQM and Data Certification services that filter out the events considered appropriate for processing [77, 78]. The datasets are available in CMS DAS and can be queried with the dataset identification strings listed in the first column [79].

Process	Dataset name	Cross section [pb]	N_{event}
$t\bar{t}h$	/ttHJetToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8_mWCutfix/[3]	0.215	2969555
$t\bar{t}W$	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/[2,5]	0.204	2700951
$t\bar{t}Z$	/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[3]	0.273	3685750
$t\bar{t}+\gamma+$ jets	/TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/[1]	3.70	4742775
single $t/\bar{t}+\gamma+$ jets	/TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8/[1]	2.97	364136
$W+\gamma\gamma+$ jets	/WGToLNuG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[3]	586	9683019
$Z/\gamma^*+\gamma$	/ZGTo2LG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[3]	131	9242330
WW	/WWTo2L2Nu_13TeV-powheg/[1]	10.5	1999091
	/WWTo2L2Nu_DoubleScattering_13TeV-pythia8/[1]	0.173	999310
	/WpWpJJ_13TeV-powheg-pythia8/[2]	0.0371	149594
WZ	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8[1]	4.43	1993205
ZZ	/ZZTo4L_13TeV-amcatnloFXFX-pythia8/[3]	1.26	6610878
WWW	/WWW_4F_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.209	210451
WWZ	/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.165	221500
WZZ	/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.0557	216284
ZZZ	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.0140	213086
single $t/\bar{t}+Z$	/tZq_ll_4f_13TeV-amcatnlo-pythia8/[3]	0.0758	3805895
$tt\bar{t}\bar{t}$	/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8/[1]	0.00910	104758
$t\bar{t}+$ jets	/TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/[1,3]	87.3	30227418
	/TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/[1]	182	61475579
	/TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/[3]	182	60202982
$Z/\gamma^*\to\ell\ell$	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	6030	35252387
	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/[4]	5770	143970947
$W+$ jets	/WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	61500	16482922
single $t/\bar{t}$	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1/[1]	3.68	622641
	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/[1]	136	67070045
	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/[1]	81	38793427
	/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/[3]	35.6	6942327
	/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/[3]	35.6	6884041

[1] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v1/MINIAODSIM
[2] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6_ext2-v1/MINIAODSIM
[3] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6_ext1-v1/MINIAODSIM
[4] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6_ext1-v2/MINIAODSIM
[5] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6_ext1-v3/MINIAODSIM

Table D.5: MC samples used in this analysis (available via DAS [79]). The full sample name is split into two parts: primary dataset name (the second column) and specific production version (given below the table). The primary dataset name includes a descriptive process name, c.o.m energy and name of the MC generator.

Observable	Loose	Fakeable	Tight
p_T	$> 7\,\text{GeV}$	$\left.\begin{array}{ll} 0.85 \cdot p_T^{\text{jet}} & \text{if } \Delta R(e,\text{jet}) < 0.4 \\ p_T & \text{otherwise} \end{array}\right\} > 10\,\text{GeV}$	$> 10\,\text{GeV}$
$\lvert\eta\rvert$	< 2.5		
$\lvert d_{xy}\rvert$	$< 0.05\,\text{cm}$		
$\lvert d_z\rvert$	$< 0.1\,\text{cm}$		
$\lvert\text{SIP}_{3D}\rvert$	< 8		
I_{mini}	< 0.4		
EGamma POG MVA	✓		
Conversion rejection	✓		
Missing hits	$\leqslant 1$	0	0
Electron cluster cuts	✗	✓ if $p_T > 30\,\text{GeV}$ ✗ otherwise	✓ if $p_T > 30\,\text{GeV}$ ✗ otherwise
$t\bar{t}h(e)$ MVA	✗	✗	> 0.75
p_T^e/p_T^{jet}	✗	> 0.3 if MVA $t\bar{t}h(e) \leqslant 0.75$ ✗ otherwise	✗
CSV of the closest jet	✗	$<$ CSVL if MVA $t\bar{t}h(e) \leqslant 0.75$ $<$ CSVM otherwise	< 0.8484

Table D.6: Electron selection criteria. Loose leptons (second column) are used to train $t\bar{t}h(e)$ MVA, fakeable leptons (third column) are used in background estimation (see section 4.3.2) and tight leptons (fourth column) are used in the event selection for the signal region (see section 4.3), as explained in section 4.2.1. Electron cluster cuts are detailed in table D.10 and lepton mini-isolation I_{mini} is discussed in appendix B.3. The symbol ✓(✗) means that the criterion is applied (is not applied). The EGamma POG MVA cut depends on $|\eta|$ and p_T of the electron as explained in table D.8. The variable p_T^{jet} refers to the transverse momentum of the jet closest to the electron. Conversion rejection cut eliminates electrons which are near oppositely charged particles. The acronyms CSVL and CSVM refer to the loose and medium WPs of the CSV b-tagging algorithm, as described in section 4.2.3.

Pseudorapidity range	$\mathcal{A}(e)$	$\mathcal{A}(\mu)$
$0.0 < \lvert\eta\rvert < 0.8$	0.1607	0.1322
$0.8 < \lvert\eta\rvert < 1.3$	0.1579	0.1137
$1.3 < \lvert\eta\rvert < 2.0$	0.1120	0.0883
$2.0 < \lvert\eta\rvert < 2.2$	0.1228	0.0865
$2.2 < \lvert\eta\rvert < 2.5$	0.2156	0.1214

Table D.7: Effective areas $\mathcal{A}$ used to correct the isolation of electrons and muons for PU effects [49, table 11]. See appendix B.3 for further discussion.

	$0 \leqslant \lvert\eta\rvert < 0.8$	$0.8 \leqslant \lvert\eta\rvert < 1.479$	$1.479 \leqslant \lvert\eta\rvert < 2.5$
$p_T < 10\,\text{GeV}$	-0.3	-0.46	-0.63
$p_T > 10\,\text{GeV}$	$\min\left\{\max\left\{T_\eta - \frac{T_\eta - L_\eta}{10} \cdot \left(\frac{p_T}{\text{GeV}} - 15\right), L_\eta\right\}, T_\eta\right\}$		

Table D.8: Cuts on the EGamma POG MVA as function of electron p_T and η. The values L_η and T_η are defined in table D.9. The formula for electrons of $p_T > 10\,\text{GeV}$ interpolates between two WPs such that looser cuts are applied for electrons of higher p_T.

	$0 \leqslant \lvert\eta\rvert < 0.8$	$0.8 \leqslant \lvert\eta\rvert < 1.479$	$1.479 \leqslant \lvert\eta\rvert < 2.5$
L_η	-0.96	-0.96	-0.95
T_η	-0.86	-0.85	-0.81

Table D.9: Definition of the symbols L_η and T_η used in table D.8.

Observable	$\lvert\eta\rvert < 0.8$	$0.8 < \lvert\eta\rvert < 1.479$	$\lvert\eta\rvert > 1.479$
$\sigma_{i\eta i\eta}$	< 0.011	< 0.011	< 0.030
H/E	< 0.10	< 0.10	< 0.07
$\lvert\Delta\eta_{\text{in}}\rvert$	< 0.01	< 0.01	< 0.008
$\lvert\Delta\phi_{\text{in}}\rvert$	< 0.04	< 0.04	< 0.07
$1/E - 1/p$	$-0.05 < 1/E - 1/p < 0.010$	$-0.05 < 1/E - 1/p < 0.010$	$-0.05 < 1/E - 1/p < 0.005$

Table D.10: Compactness cuts on electron energy deposited in the ECAL that are applied as part of electron selection defined table D.6. The variable $\sigma_{i\eta i\eta}$ denotes the cluster width in η direction; H/E is the ratio of the energy in the HCAL tower directly behind ECAL's SuperCluster (H) to the SuperCluster energy (E); $\Delta\eta_{\text{in}}$ and $\Delta\phi_{\text{in}}$ are the differences in η and ϕ direction between energy-weighted SuperCluster position and the track extrapolated to the ECAL; p is the reconstructed track momentum at the point closest to the PV ($1/E - 1/p$ is therefore the deviation of the SuperCluster energy from measured momentum). These variables are also used in the electron identification algorithms. The cuts reassure that the electron track is correctly linked to its associated ECAL SuperCluster, and that the electron deposits most of its energy in the ECAL, assuming little to no *bremsstrahlung*. The cuts also exploit the fact that the cone size of electromagnetic showers tends to be narrower than that of hadronic showers. All momenta and energies are given in units of GeV.

Observable	Loose	Fakeable	Tight
p_T	$> 5\,\text{GeV}$	$> 10\,\text{GeV}$	$> 10\,\text{GeV}$
$\lvert\eta\rvert$	< 2.4		
$\lvert d_{xy}\rvert$	$< 0.05\,\text{cm}$		
$\lvert d_z\rvert$	$< 0.1\,\text{cm}$		
$\lvert\text{SIP}_{3\text{D}}\rvert$	< 8		
I_{mini}	< 0.4		
Loose PF muon	✓		
Medium PF muon	✗	✗	✓
$t\bar{t}h(\mu)$ MVA	✗	✗	> 0.75
$p_T^\mu / p_T^{\text{jet}}$	✗	> 0.3 if MVA $t\bar{t}h(\mu) \leqslant 0.75$ ✗ otherwise	✗
CSV of the closest jet	✗	$<$ CSVL if MVA $t\bar{t}h(\mu) \leqslant 0.75$ $<$ CSVM otherwise	< 0.8484

Table D.11: Muon selection criteria. The symbol ✓(✗) means that the cut is applied (is not applied). The acronyms CSVL and CSVM refer to the loose and medium WPs of the CSV b-tagging algorithm, as described in section 4.2.3.

Process		Branching ratio	
$\pi^+ \to \mu^+\nu_\mu$		99.98%	
$\pi^0 \to 2\gamma(\to e^+e^-\gamma)$		98.82% (1.17%)	
$K^+ \to \mu^+\nu_\mu$		63.56%	
$\tau^- \to h^-\nu_\tau$	$\tau^- \to \pi^-\nu_\tau$	10.82%	11.53%
	$\tau^- \to K^-\nu_\tau$	0.696%	
$\tau^- \to h^-\nu_\tau + 1\pi^0$	$\tau^- \to \pi^-\pi^0\nu_\tau$	25.49%	25.93%
	$\tau^- \to K^-\pi^0\nu_\tau$	0.433%	
$\tau^- \to h^-\nu_\tau + 2\pi^0$	$\tau^- \to \pi^-2\pi^0\nu_\tau$	9.26%	9.32%
	$\tau^- \to K^-2\pi^0\nu_\tau$	0.065%	
$\tau^- \to h^-h^+h^-\nu_\tau$	$\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	9.31%	9.80%
	$\tau^- \to \pi^-\pi^+K^-\nu_\tau$	0.345%	
	$\tau^- \to \pi^-K^+K^-\nu_\tau$	0.144%	

Table D.12: Branching ratios for different hadronic τ decay modes [24]. The symbol $h^\pm$ stands for either a charged pion ($\pi^\pm$) or kaon ($K^\pm$). Decay modes containing one $h^\pm$ are referred to as „1-prong", whereas the decay modes with three $h^\pm$ are called „3-prong" hadronic τ decays.

Observable	Fakeable	Tight
p_T	$> 20\,\text{GeV}$	
$\lvert\eta\rvert$	< 2.3	
$\lvert d_z\rvert$	$< 0.2\,\text{cm}$	
Tau POG MVA ($R = 0.3$)	loose WP	medium WP

Table D.13: Selection criteria for τ_h.

Observable	Loose	Medium
p_T	$> 25\,\text{GeV}$	
$\lvert\eta\rvert$	< 2.4	
CSV WP	$>$ CSVL $= 0.5426$	$>$ CSVM $= 0.8484$

Table D.14: Selection criteria for b-jets.

Trigger type	HLT path	Primary dataset	Event category
$1e$	HLT_Ele25_WPTight_Gsf HLT_Ele27_WPTight_Gsf HLT_Ele25_eta2p1_WPTight_Gsf HLT_Ele27_eta2p1_WPLoose_Gsf	SingleElectron[5]	$3e$, $2e1\mu$, $1e2\mu$
1μ	HLT_IsoMu22 HLT_IsoTkMu22 HLT_IsoMu22_eta2p1 HLT_IsoTkMu22_eta2p1 HLT_IsoMu24 HLT_IsoTkMu24	SingleMuon[4]	$2e1\mu$, $1e2\mu$, 3μ,
$2e$	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	DoubleEG[3]	$3e$, $2e1\mu$
$3e$	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL		$3e$
2μ	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ	DoubleMuon[1]	$1e2\mu$, 3μ
3μ	HLT_TripleMu_12_10_5		3μ
$1e1\mu$	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL	MuonEG[2]	$2e1\mu$, $1e2\mu$
$2e1\mu$	HLT_Mu8_DiEle12_CaloIdL_TrackIdL		$2e1\mu$
$1e2\mu$	HLT_DiMu9_Ele9_CaloIdL_TrackIdL		$1e2\mu$

Table D.15: HLT trigger paths and associated primary datasets (PDs). Since the PDs have been recorded with different triggers, we have prioritized the PDs to avoid double-counting of events. The priorities are indicated with superscripts after the PD name. The HLT names indicate which selection criteria a given event has to satisfy in order to pass the trigger. The names consists of lepton flavors, their multiplicities, lower p_T bounds and optionally track or calorimeter parameters. Events selected in the analysis have to pass at least one of those triggers. The last column specifies the event categories on which the triggers bits are applied. For instance, if an event contains 2 electrons and 1 muon on the level of the offline event selection, we require that at least one of the triggers $2e1\mu$, $1e1\mu$, $2e$, $1e$ or 1μ has fired. If none of these triggers has fired, then the event is excluded from further analysis.

	Systematics	Type	Applies to	Uncertainty
Experimental	Luminosity	N	MC	2.6%
	Trigger efficiency	N	MC	3%
	τ_h identification	N	MC	5%
	e, μ identification	N	MC	e 2%, μ 3%
	jet-to-τ_h FR	SN	MC	-
	fakeable-to-tight FR of e, μ	SN	fake background	-
	b-tagging	SN	MC	-
	JEC	SN	MC	-
	τ_h energy scale	SN	MC	3%
Theoretical	QCD scale	N	$t\bar{t}h$, $t\bar{t}W$, $t\bar{t}Z$ MC	$^{+5.8}_{-9.2}$%/$^{+12.9}_{-11.5}$%/$^{+9.6}_{-11.3}$ %
	PDF	N	$t\bar{t}h/t\bar{t}W/t\bar{t}Z$ MC	±3.0%/±2.0%/±2.8%
	α_S	N	$t\bar{t}h/t\bar{t}W/t\bar{t}Z$ MC	±2.0%/±2.7%/±2.8%
	μ_F and μ_R variation	S	$t\bar{t}h/t\bar{t}W/t\bar{t}Z$ MC	-
	background yield	N	rare SM & WZ	50%

Table D.16: List of systematic uncertainties. The second column indicates whether a given systematic uncertainty affects the shape (S) or the normalization (N) of the discriminating variable that is used for the signal extraction (c.f. section 4.4.2). The third column lists the signal and/or background processes on which the systematics is applied to and the fourth column shows the resulting uncertainty on the event yield. The systematic uncertainties that change only the event yields are added in quadrature after the event selection, so there is no need to re-run the analysis to estimate their impact on the event yields.

$\cos\theta_1$	ϕ_1	$\cos\theta_2$	ϕ_2	z_1	$\phi^{\text{inv}}(\tau_h, \nu)$	$\phi^{\text{inv}}(\ell, 2\nu)$	m^2_{inv}
$(-1, 1)$	$(-\pi, \pi)$	$(-1, 1)$	$(-\pi, \pi)$	$(0, 1)$	$(-\pi, \pi)$	$(-\pi, \pi)$	$(0, m^2_\tau)$

Table D.17: Integration variables and limits used to evaluate the integral (4.9) for the MEM. The physical meaning of the integration variables is explained in section 4.4.3.

$\cos\theta_1$	ϕ_1	$\cos\theta_2$	ϕ_2	z_1	$\phi^{\text{inv}}(\tau_h, \nu)$	$\phi^{\text{inv}}(\ell, 2\nu)$	m^2_{inv}
1.00000	0.07673	0.07842	0.00833	0.00261	-0.07397	-0.05446	0.00148
0.07673	1.00000	-0.03842	-0.06090	0.00581	-0.02888	-0.03236	-0.01733
0.07842	-0.03842	1.00000	0.03568	-0.00056	0.02716	-0.06968	0.00498
0.00833	-0.06090	0.03568	1.00000	0.07607	-0.02645	0.03064	0.07589
0.00261	0.00581	-0.00056	0.07607	1.00000	-0.04036	0.04325	0.01776
-0.07397	-0.02888	0.02716	-0.02645	-0.04036	1.00000	-0.03152	0.05478
-0.05446	-0.03236	-0.06968	0.03064	0.04325	-0.03152	1.00000	-0.02970
0.00148	-0.01733	0.00498	0.07589	0.01776	0.05478	-0.02970	1.00000

Table D.18: Correlation matrix between the integration variables used in the MEM. The numbers are computed from generator-level information of simulated $t\bar{t}h$ signal events that selected in the SR. All reconstructed objects in the selected events are required to be matched to the corresponding generator-level objects.

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Karl Ehatäht (sünnikuupäev: 09.09.1991),

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

   „Search for $t\bar{t}h$ production in $3\ell + 1\tau_h$ in final state",

   mille juhendajad on Christian Veelken ja Luca Marzola,

   1.1 reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

   1.2 üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 26.05.2017